Part 2

Analysing Risks with Financial Models: Sensitivity Analysis, Scenario Analysis, Break-even Analysis, Time Series and Monte Carlo Simulation

Chapter 14: Risk Assessment: the Centrepiece of All Valuation, Contacting and Credit Issues in Finance and the Basis for Making Judgments in your Life

One of the tasks required to manage any business is the honest and relatively boring job of accounting for its historic financial results through keeping track of profit and asset value in financial statements. A perhaps less noble and much more challenging task faced by all but the very simplest of business endeavors and by those who want to invest in the business is to make implicit or explicit forecasts of future cash flow. The problem with making such a forecast is that they by definition encompass an unknown and uncertain future. With the possible exception of forecasting nominal cash flow earned on risk free bonds, all cash flow forecasts will turn out to be wrong. Financial modelling techniques described in the last part generally covered how to make a financial model but did not address how to quantify potential uncertainty in assumptions for key variables that drive projected revenues, operating expenses and capital expenditures. This second part of the text moves to the more difficult and interesting question of how to assess the uncertainty associated with future economic variables such as price, demand, expense structure, cost of new capacity and other items. The amount by which implicit or explicit forecasts of cash flow could be different from realized results (rather than beta, VAR or other elaborate statistics) forms the framework for thinking about risk in this part of the text.

Assessing the risk associated with potential variation in future cash flows is at the heart of just about every issue in finance and many other professions. When Winston Churchill stated: “true genius resides in the capacity for evaluation of uncertain, hazardous, and conflicting information” he was probably not talking about adding risk to financial models, but he could have been. The process of evaluating uncertainty associated with how much cash flow forecasts will differ from expected levels may involve estimation of how much estimated growth rates will be wrong; what will be the possible change in interest rates from base case forecasts; how much actual product prices and costs will be different from predicted prices and costs; what will be the change in public attitudes to various products; or what sudden changes in political events could occur. This somewhat ambiguous concept of risk defined as the potential variation in outputs of a financial model is not necessarily consistent with the attempt by many finance academics who suggest that risk can be captured in a single number such as the beta of a stock, the volatility of prices or the value at risk of an investment portfolio.

Attempting to measure the amount by which a financial model can be wrong is by no means a new subject in finance and over the past fifty years Nobel Prizes have been awarded to economists who have developed various “revolutionary” mathematical or statistical approaches to measuring risk ranging from beta to probability of default to value at risk. Notwithstanding the elegant formulas and complex statistics, application of new and seemingly promising innovative mathematical methods to risk assessment have generally turned out to be frustrating in practice. After the global financial crisis of

2008, more bankers, stock analysts, financial managers and consultants would rather perform a scenario analysis than attempt to represent a company as a mathematical equation with a probability distribution. As more and more complex mathematical risk assessment tools are developed, conflicts between whether risk should be evaluated using mathematical methods or whether risk should be measured by business judgment has become a issue that is lively debated in finance.

In describing mechanics of measuring risk, this part of the book begins by reviewing risk assessment techniques that rely on business savvy and economic judgment – qualitative risk matrices, sensitivity analysis, break-even analysis, scenario analysis and tornado diagrams. As with the earlier explanation of financial modelling concepts, fundamental risk analysis is described using step by step explanations of how to add various features to a financial model when assessing risk. After addressing risk assessments that require judgment with respect to what can happen to a variable, the remainder of this chapter describes how to create stochastic time series models statistics to measure risk on a mathematical basis. The discussion of mathematical risk analysis demonstrates how one can use a combination of mathematics and judgment to create probability distributions of various assumptions in a financial. To develop mathematical models and probability distributions, a statistical tool kit of statistical parameters is introduced that includes volatility (the dispersion in prices), mean reversion (the speed at which prices come back to long-run average levels), correlations with other prices (such as the correlation between natural gas and oil prices), lower and upper price boundaries on price movements, price jumps, price trends and long-run equilibrium prices.

Six Alternative Ways to Assess the Risk of a Company, a Project or a Contract

These days, corporations hire people with finance, mathematics, physics or economic degrees and give them a title of “risk manager.” Presumably, in one way or another, risk managers are supposed to evaluate how various decisions affect the risk and value of equity and debt investors. The somewhat mysterious job title seems to encompass everything from a junior analyst computing historic financial ratios for credit reviews to a person with a PhD in physics who develops value at risk statistics derived from complex mathematical equations. In practice, quantification of risk can encompass a whole lot of different analyses ranging from simply graphing an economic variable (prices, costs, capital expenditures etc.) and observing how that variable affects cash flows and value (sensitivity analysis), to computing the probability distribution of cash flows using time series equations and Monte Carlo simulation.

The different sorts of risk analysis can be classified is according to their mathematical complexity. At one end of the risk assessment scale is a purely qualitative list of risks along with their importance and possible mitigation. Other approaches along the mathematical complexity spectrum move from sensitivity analysis to break-even analysis to scenario analysis to tornado diagrams and finally to Monte Carlo simulation. The table below lists six different approaches to risk analysis along with a brief description of what sort of analysis is involved in each technique.

Table 1

Risk Matrix	Sensitivity Analysis	Break-Even Analysis	Scenario Analysis	Tornado Diagram	Monte Carlo Simulation
List Economic and Financial Factors that can Influence the Financial Performance of an Investment and Determine whether the Risks are Mitigated through Contract Provisions and/or Hedging. For Variables that are not Mitigated, use Adjacent Risk Analysis Techniques	Choose an Important Economic Variable that is not Mitigated and make a Graph of how the Variable Effects the Outputs of Financial Variables Related to Valuation such as Net Present Value and IRR (The IRR can be the IRR on Equity, IRR on Debt or the Overall IRR of the Project Free Cash Flows)	Make a Table that Computes Outputs of Financial Variable Along Side of an Economic Variable and Determine when the Financial Variable Becomes Unacceptable. Once the Break-even Level is Established, one can Evaluate the Likelihood of the Economic Variable becoming the break-even Level.	Evaluate a Series of Different Output Variables Given a Set of Different Input Variables. For example, a Downside Case can be Established then Different Levels of Debt can be Assessed in the Downside Case. If the Downside Case cannot Support the Level of Debt, then Alternative Debt Levels are Recommened	The Tornado Diagram evaluates which variables have the most Significant Effect on the Financial Output of a Model Related to Valuation and which Variables have a Relatively Insignificant Impact. This tool can be used for Due Diligence Analysis or to Demonstrated the Relative Upside Potential and Downside Risk	Simulation Produces a Distribution of Returns and can Provide the Likelihood of Valuation Variables such as IRR and NPV Falling within a Certain Range. The Simulation Depends on Time Series Equations that are Derived from Parameters such as Volatility, Mean Reversion, Correlation and Price Boundaries

The above table puts judgmental approaches to risk assessment in the left columns and moves to the more mathematical methods on the columns listed at the right. A lot of the discussion in this chapter addresses Monte Carlo simulation -- the rightmost column of the table. The lack of business judgment required in this approach does not at all imply that Monte Carlo models produce a better assessment of risk than careful business judgment regarding how a particular variable can move in the future. Very intelligent people who have made a large personal investment in the study of complex mathematical approaches to risk assessment have a natural desire to apply their knowledge in practice. However most of the valuation mistakes discussed in Chapter one such as Eurotunnel and AES Drax would surely not have been solved through the more use of elaborate simulation and time series models. Indeed, some business valuation mistakes have been aggravated by inappropriately using historical data in applying complex models.

A valuation nightmare that illustrates conflicts between different ways to assess risk is the case of Long-term Capital Management (LTCM) that was a big story in the late 1990's. LTCM was a hedge fund created in part by Nobel Laureates Myron Scholes and Bob Merton that, when it failed, almost brought the entire financial system to its knees and required a federal government bailout (although the problems seem like small change compared to Lehman Brothers and other problems of 2008.)[1] The company was created to make money by developing sophisticated mathematical models that evaluated statistical relationships and supposedly hedged many risks. As risks were measured through complex mathematical techniques, the heart of the LTCM case is the question of whether mathematical models and statistics can replace business judgment when gauging risk.

In making investments on the basis of statistical analysis and mathematical models, the LTCM case contained an explosive mixture of the valuation errors discussed in Chapter one that includes:

- Assumption that historic statistical relationships can be used to predict the future and be re-established after a period of time. LTCM failed when many economic variables did not act as they had in the past and sudden non-linear shocks occurred without historic precedent;

1 Cite "When Genius Failed" and the Trillion Dollar Bet

- Non-transparency in presenting financial information. Both investors and lending banks (the largest Wall Street Banks) were not allowed to see how trades were made and what models were used;

- Belief that staff employed by the fund was smart enough to continually beat the market and earn returns well above the cost of capital in extremely competitive financial markets where risk adjusted returns should not be very high. This notion that the fund could continually earn high returns ultimately caused the fund to take higher and higher risks (like a lot of other companies) when returns began to fall;

- Investment strategies that were derived from highly complex mathematical models that supposedly discovered new ways to make money without being fully vetted. If LTCM could really make money by studying mathematical relationships (anathema to believers in efficient markets such as Scholes and Merton), then others would be able to copy the formulas and eliminate the profits;

- Faith in the reputation of other people -- Merton, Scholes and a famous bond trader named Bruce Meriwether -- without independently assessing whether the fundamental business concepts were viable; and,

- Investments that were made without verifying the underlying economics with simple back of the envelope checks. For example, the fund invested in Russian bonds by studying mathematical relationships rather than by visiting the country and seeing that military officers and secretaries were not being paid (in 1998 when Russia defaulted on its bonds, the country was arguably in a worse financial state than when the Soviet Union collapsed).

The LTCM case raises a fundamental question of whether risk analysis can be reduced to a series of mathematical equations or whether risk must ultimately be assessed through business savvy and judgment. LTCM’s dramatic demise demonstrated that complex quantitative techniques for risk measurement cannot be used without any supplemental business judgment. The healthy tension between relying on mathematical equations and using judgment raised in the LTCM case provides a backdrop in considering how best to assess different approaches to risk assessment.

Direct Risk Assessment to see what Happens to Cash Flow and Financial Ratios Versus use of Beta and WACC to Indirectly Measure Risk

None of the methods shown in Table 1 mentions the single parameter that finance theory implies can measure the relative risk of a stock, i.e. beta. Finance courses and textbooks suggest that one has to go no further than to measure beta from stock prices (which converts into a weighted cost of capital number) and then use this risk adjusted discount rate along with expected cash flows to measure value – no scenario analysis, no Monte Carlo simulation or any other direct measurement of risk is necessary. As beta supposedly contains all relevant information about risk, all of the techniques in Table 1 would not be necessary. Risk analysis methods presented in this chapter are premised on the notion that risks cannot be magically be stuffed into one cost of capital number and involve a search for risk analysis methods other than traditional discounted cash flow and weighted average cost of capital analysis. To contrast risk analysis using beta and weighted cost of capital with other approaches, consider the

following two different hypothetical presentations made by different advisors to a company thinking about acquiring a company:

- Method 1: Valuation from Financial Theory

  The first advisor values the target company by carefully applying techniques taught in business school including computation of free cash flow, measurement of the weighted average cost of capital and use of on-going growth rates after the explicit forecast period. He makes elegant calculations of the cost of equity capital from statistical analysis of historic stock prices and applies market weights to the capital structure. This method results in valuation of the target company on a standalone basis that is then adjusted for the value of synergies generated from the merger. Many people looking at the DCF analysis will immediately recognise that two of the most subjective assumptions in the valuation – the weighted average cost of capital and the growth rate – can lead to a very wide variation in value as illustrated in the table below.

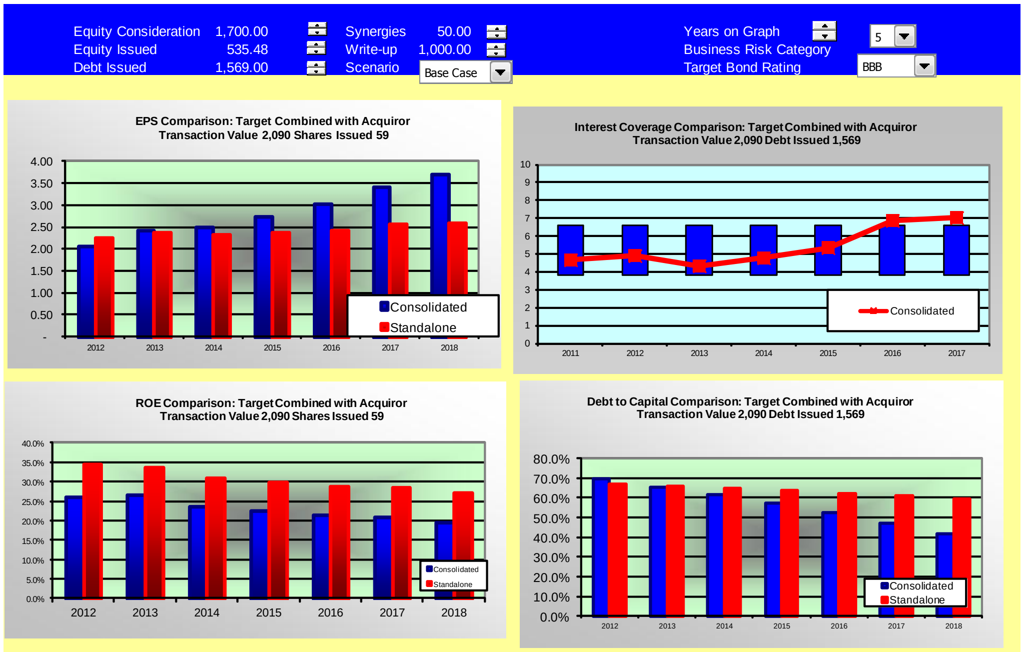
Range in Implied EV/EBITDA from Applying the Discounted Cash Flow Model

Weighted Averge Cost of Capital

	$7.16	7%	8%	9%	10%
	0%	8.34	7.26	6.42	5.75
Terminal	1%	9.38	8.00	6.97	6.17
Growth	2%	10.82	8.99	7.68	6.70
Rate	3%	12.99	10.37	8.62	7.37
	4%	16.61	12.44	9.94	8.27

- Method 2: Accretion and Dilution from Financial Models

  The second advisor uses another method to asses risk and value of the merger. He creates two financial models: one for the acquiring company without a merger and another for the combined acquiring company and the target company using the integrated merger model techniques discussed in the last chapter. Then, after synergies are included in the analysis, he makes a presentation of whether earnings per share are higher in the combined case (accretion), or whether earnings per share decline (dilution). He also makes an assessment of whether the company can maintain its investment grade bond rating given the proposed financing in the merger. This integrated analysis accounts for the actual proposed financing of the transaction, the effect of the transaction on interest costs and stock prices as well as the fees paid for advisory services. The maximum value to be paid for the target company is the purchase price number that just avoids dilution in earnings per share. However in the presentation made by this second investment banker, there is no WACC to be computed, no terminal value to be estimated and the investment banker can apply similar assumptions for the acquiring company and the target company.



Knowing that valuations using the DCF technique can be so easily manipulated by making small changes in the WACC or the terminal growth rate – two variables which are very difficult if not impossible to measure – the company considering the merger most probably would pay more attention to the technique presented by the second advisor. Then, rather than pretending that risk of the making the acquisition can really all be incorporated in the WACC, the decision maker could ask for a series of different scenarios and break-even analyses that test how sensitive the accretion or dilution estimates are to various key variables with drive the forecasts. For example, he may evaluate the break-even point for the growth rate before which the merger becomes dilutive. Here, the financial officer is directly assessing risk rather than assuming it can be stuffed into the beta parameter. In a similar vein, decision makers assessing the risk of a project financed investment or a leveraged acquisition can directly assess risks to the equity IRR falling below a certain level and the coverage of debt service using analogous techniques. These techniques to directly measure risk rather than adjusting the cost of capital are the subject of the remainder of this chapter (the subject of cost of capital and beta are covered in the next chapter.)

Chapter 15: Approaching Risk Analysis through making Judgements that Categorise Risks, Estimate Their Importance and List Contracts, Hedges and Insurance that can Mitigate Risks

This section describes practical issues that arise in implementing the first five approaches to risk assessment listed in Table 1. After describing how the various approaches can be applied to analyse risk, a step by step process of how to implement the technique is presented. As with the discussion in Chapter two, you can skip the detailed instructions without missing the general idea of how the different methods work.

The First Step in Risk Assessment: Listing Risk issues and a Risk Allocation Matrix that Includes the Importance of Each Risk, Mitigation Measures and the Magnitude of Residual Risks

A useful first step in directly assessing risk is to categorise, describe and weight risks in some structured manner rather than diving into any quantitative analysis. There is no mathematical analysis whatsoever in this risk assessment technique. Instead, the process should force you to think about, discuss and understand a variety of different risks; how the risks can be mitigated by contracts; and, how those risks could potentially affect investment returns. Qualitative analysis may involve preparing a memo on key risk issues, categorizing various risks into a matrix that describes risks or some other approach. In developing a list of risks, the risk matrix often ends up to be rather boring and mechanical where construction delay risks are mitigated by liquidated damages; interest rate changes are mitigated by interest rate swaps; operation and maintenance expenses are mitigated by a fixed price contract. One of the interesting elements in preparing the risk matrix is to be able to imagine seemingly minor risks that can explode into big problems. Further, when thinking about categorising and then mitigating risks, one of the most useful things is to put some kind of weighting on the different risks and to realise that even when a risk is seemingly mitigated, the risk is transferred to a contract that may not be sustained meaning that the risk is not really completely eliminated.

Once the risks are defined and described, one can consider whether the some of the risks can be mitigated through one of three techniques -- insurance, contracts and or hedges. Mitigation of risks using one of these three methods generally involves some kind of explicit or implicit cost and an important valuation skill is the ability to consider tradeoffs between the benefits and costs of mitigating different risks. The problem with lack of transparency in reporting is that analysts cannot even identify the risks. To see how risk matrices can be useful in evaluating the cost and benefit of mitigation strategies you could imagine applying the process in health. Health professionals tend to only focus on attempting to minimise a risk without evaluating the cost of the risk mitigation in terms of lifestyle and monetary terms.

If all risks could somehow be perfectly mitigated (if you did this in your life it would be unbearably boring), then the investment would be risk free and values could be derived by simply discounting cash flows at the risk free interest rate. In most investments, some risks will not be mitigated by insurance, contracts or hedging (even if the risks can seemingly be mitigated by contracts, contracts will likely be broken for large risks). For those risks that are not mitigated, one can apply one or more of the subsequent risk assessment techniques such as break-even analysis to evaluate the magnitude of the risk and whether the risk is acceptable for lenders and/or equity providers. Consider an extreme case where all of the risks are mitigated except for changes in the interest rate. Once the risk matrix demonstrates this fact that all risks are mitigated, one can determine how high interest rates would have to move before the investment cannot pay back its debt (from a lender perspective) or how high interest rates would have to rise before returns to equity holders fall below the risk free rate.

The table below illustrates the types of items that could be included in a risk matrix. Items to include in the table include the likelihood and the impact of a risk (a low probability risk with high consequences and a relatively high probability risk with relatively low consequences); the technique for mitigation of the risk using contracts; problems with the risk mitigation such as a counterparty default on contracts; the relative importance of residual risks not mitigated; and the analysis used to evaluate the residual risks. The column on the left lists selected cases in which un-mitigated risks caused problems with the performance of investments.

Table 2

Risk Category	Description	Mitigation	Analysis of Un-Mitigated Risk	Example
Construction Phase				
Construction Over-run				
Material Prices		Contracts/Hedges	Break-even effect on IRR/Debt	Petro Chemical Plants
Cost Plus Provisions		Contracts	Cost of Comparable Projects	Eurotunnel/Eurodisney
Exchange Rates		Hedges	Break-even/Scenario/Monte Carlo	Petrozuata
Force Majure		Insurance		
Construction Delay				
Primary Project		Contracts-Liquidated Damage	Break-even effect on IRR/Debt	A380/US Nuclear Plants
Associated Projects		Contracts	Break-even effect on IRR/Debt	Eurotunnel
Technical Failure				
Resource Quantity		Resource Studies/Tests	Reserve Report	
Plant Efficiency		Contracts-Liquidated Damage	Engineering Report	Alstrom Combined Cycle
Operation Phase				
Price		Contracts/Hedges/Options	Break-even/Scenario/Monte Carlo	Argentina Merchant Plants
Volumes		Contracts	Break-even/Tomdao Diagram	Uncongested Toll Roads
Cost		Contracts/Hedges	Break-even/Tomdao Diagram	MCV Cogeneration Plant
Royalties		Contracts/Political Insurance	Break-even/Scenario/Monte Carlo	Petrozuata
Political				
Contract Abrogation		Political Insurance	Price Analysis/Off-taker Credit	Enron Dabhol/AES Drax
Nationalization		Political Insurance	Option Price Analysis	Petrozuata
Currency Convertibiltiy		Political Insurance		Zimbabuae Mines
Political Unstability		Political Insurance		Gaza Power Plant
Financial				
Interest Rate Changes		Hedges/Options	Break-even/Scenario/Monte Carlo	Pub Service New Hampshire
Inflation Rate Changes		Contracts	Break-even effect on IRR/Debt	PT Pation Indonesia
Re-financing Risk		Hedges/Options	Break-even/Scenario/Monte Carlo	Sub-prime Crisis

Chapter 16: Adding Sensitivity Analysis to Financial Models through Creating Dynamic Graphs that Show How Key Inputs Affect Various Outputs

One of the simplest and most effective ways to analyse risk is through simply changing one of the inputs to the financial model that is not mitigated with a contract and gauging the effect on key output variables such as the value of the company, the IRR or the debt service coverage ratio. This involves making an effective presentation as to what happens to some measure of value when the level of a key variable changes. High-powered analysts who compute statistics such as value at risk, probability of default and complex option price premiums, may scoff at sensitivity analysis as being overly simplistic. But consider a variable is very difficult to predict such as the level of traffic that will occur between England and France in the Eurotunnel, demand growth in the Philippines, the level of housing prices or electricity prices after a change in the structure of a market – all of which were discussed in the Chapter one valuation nightmares. Given the difficulty in creating an equation that can predict any of these variables, an effective picture of what happens to the value of the investment when the variable changes may be the best that one can realistically accomplish. For example, in assessing the value of collateral debt obligations that consisted of sub-prime mortgages, it surely would have been very useful to have a picture of what would happen to the value of the investments under alternative housing demand and supply assumptions. Alternatively, consider the case of Constellation Energy discussed in the Chapter one. Management presented value at risk statistics and maintained that the purpose of its trading

activities was primarily to hedge other parts of the business (recall that management encouraged equity analysts simply to trend cash flows into the future.) This risk analysis frustrated investors and provided little useful information. With hindsight, a much more helpful presentation would have been a simple sensitivity analysis that illustrated what happens to Constellation cash flow and earnings under a wide range of energy price paths.

The figure below illustrates sensitivity analysis in the case of evaluating acquisition of an oil company, where acquisition uses senior debt, subordinated debt, preferred stock and common equity. A series of controls above the dashboard allows one to examine how projected cash flow is distributed using alternative financing structures and different operating assumptions. For example, through pushing the oil price up or down, one can observe the length time for the repayment of a loan, the rate of return earned on different securities and the value of the company and the point at which the a given rate of return cannot be achieved. For a senior debt, subordinated debt or equity investor, one could imagine decision makers varying the oil price and evaluating what happens to their cash flow. Rather than wasting a lot of time on forecasting oil prices, the investor could begin with his desired return and sees what level of oil price is required to attain the desired cash flow and the target rates of return. The picture is intended to demonstrate that sensitivity analysis is a lot about presentation and creatively summarizing key input drivers along with effective measures of value.

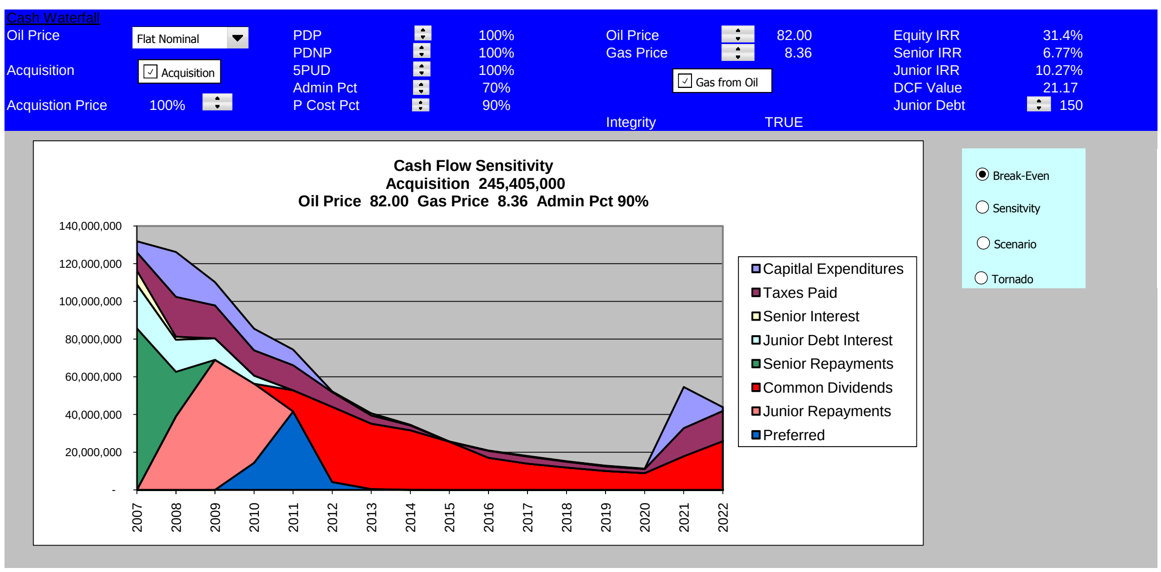


Figure 1

Converting Periodic Data into Annual, Semi-Annual or Quarterly Data for Making Things Easier to Read and Presenting Sensitivity Analysis

To make an effective presentation of sensitivity analysis, it is generally better to graph data on an annual basis rather than using detailed monthly or quarterly data. If a model is computed on a periodic basis, it is simple to annualize periodic cash flows for effective presentation and output graphs once the fiscal year has been established. It is also relatively easy to present balance accounts that do not accumulate over the year such as the closing balance of plant or debt. Annualizing can be accomplished by listing the year below the date and then using the SUMIF function (or in excel 2007

and later the SUMIFS function) to sum the periodic cash flows within the year. To demonstrate the process of creating annual data from periodic data, the following step by step process walks through the process and allows one to very quickly convert periodic cash statements to annual statements.

Step 1: Compute the a time variable (such as a year) and switch variable for the end of the fiscal year

The SUMIF or the AVERAGEIF functions compute the sum or the average across a larger series of rows or columns that will be summarised. The simplest example is if there are twelve months of a year the should be summed. To find the annual sum across many years of monthly data, you must know the year of each year that is to be summed. This is named the range in the SUMIF or the SUMIFS functions.

If you are summing balance sheet accounts, you need to identify the single period at the end of the period that will be summed. If the month of the ending date of a model is equal to the fiscal month, (e.g. June) then the fiscal month switch is TRUE.

Month(Ending Date) = End of Fiscal Year Month - 1

Step 2: To make a summary that can be used for graphs, add a new sheet that will contain annual analysis and create a simple year variable by starting with the first fiscal year and incrementing by one.

Step 3: Copy selected titles of cash flow, balance sheet and other items that you would like to present on an annual basis to the annual sheet.

Step 4: Apply the SUMIF function or the SUMIFS function for cash flow items. The SUMIF function accepts three variables -- a range; a criteria, and a sum range. The SUMIF function accepts a range from a sheet or a part of the model with detail time periods that can be summarised into aggregate periods. The range is often the year from the model that is computed on a monthly, quarterly or semi-annual basis. If a monthly model is used, then there are twelve years that have the same number. The second number required in the SUMIF function is a single criteria that will be used to summarise the data. If the sum of a data series from a monthly series is used and the sum is for the year 2016, then the year 2016 is the criteria variables. The typical way to do incorporate criteria variables is to type in a series of years -- one by one -- that correspond to the larger group of years in the detailed model. The third input is the sum range or the average range that is the range that will be summed or averaged.

An effective way to apply the SUMIF function is through clicking on the entire row or column for the year range and the sum range in the page with the detailed periodic data. The idea of using an entire row (by pressing the SHIFT and CNTL key) makes the process far easier. First click on the entire row of years in the detailed sheet and press the F4 short cut key to fix the test range name. Next, refer to year entered for the annualized report in step 2 above and use a relative reference to lock-in the row number but not the column name (press the F4 key twice). Finally, refer to the range to be summed by clicking on the entire row in the detailed sheet without locking in the row number or the column number. This process will work for the cash flow items to be accumulated over a period. In the example below row 12 in the detailed sheet with periodic flows contains the year that is used for the test. Row 5 of the annual sheet contains a single number for the row and the amount annualized is in row 20 of the detailed sheet. Again, the key to making this process easy is to use the entire rows in the detailed sheet.

Annualized Amount = SUMIF(detail!$12:$12,annal:C$5,detail!20:20)

Step 4: Sometimes you would like to present items such as debt or production capacity on the annual page that do not accumulate. To this you can use a TRUE/FALSE switch to identify codes that are accumulated over the year and the accounts in which balances are taken only from the column in the detailed sheet at the end of the fiscal year period. Then you can accumulate the amounts only for the final period in the year by using a similar SUMIF statement where the test range is a combination of both the number that is the year number and a switch variable that is true only for each period at the end of the fiscal year. You can also use the SUMIFS function that allows you to enter two criteria – one for the year number and one for the code for the end of the year defined above. The only items are summed when the fiscal year data used in the criteria of the SUMIF formula is the same as the date in the periodic section (sheet of the model.) To make this work, shade the entire row of the end date in the periodic section of the model (click on the row number) and then fix the references (press the F4 key). Finally, refer to the range to be summed by clicking on the entire row without locking in the row number or the column number as with the accounts to accumulate. In the example below, row number 21 is summed in the detailed model and row number 13 includes a switch variable that is true only at for the last period in each year (e.g. December in a monthly model):

Annualized Amount = SUMIFS(detail!21:21,detail!$12:$12,annal:C$5,detail$13:$13,TRUE)

Step 6: One of the difficulties with the process above is that when you want to pull a variable from the detailed model (that is not in the same row number order) you have to manually adjust the SUMIF formula by referring to particular row in the detailed sheet. Instead of manually adjusting the last part of the SUMIF formula, a better method is simply to enter the row number from the detailed model, insert that row number in a column in the annual sheet and then have the formula work automatically. To create a very flexible process where you do not have to change the row number in the SUMIF formulas, you can use the INDIRECT function in excel that refers to a flexible row input using the ROW function as well as a sheet name (that must be created using an excel function). Using the indirect function, you can create a range name that is used in with the SUMIF function and you can use the ROW function to select any row in the detailed sheet.

The trick to using the INDIRECT function is to create a range name that includes the sheet name. To create a range name you can use the & operator and enter the relevant symbols. For example if the detailed sheet is named detail and you would like to create a range name for row number 20 where the 20 is from a cell reference D15 (computed with the ROW function), the following steps could be used to convert the detailed data to annual data:

Step 1: Use the Row function to enter the selected row number in the annual sheet

Row number = ROW(detail!:A20)

Step 2: Unlike for a row or column, there is no function in excel to find the sheet name. Therefore you can create your own function as was the case for many examples in the modelling chapter. To create a function that can find the name of the detail sheet you can write a simple function shown below. This function, which is very helpful in a variety of circumstances because of the difficulty in finding the sheet name is simple to write once you know a rather obtuse extension called .Parent.Name:

```
Function sheet_name(cell_reference)
        sheet_name = cell_reference.Parent.Name
End Function
```

Step 3: Use the sheet name and the row number found from above to create a cell reference for the INDIRECT function by combining the cells into one range name. To make a range name from the sheet name and the row number you need to add quotation marks and codes that are used in a range name as illustrated below:

Range Name = "'" & sheet name "'!" & row number & ":" & row number

If you have a lot of instances of making a range name, you can create a function that puts in all of the painful quotation marks.

Step 4: Once the range name is defined the INDIRECT function can be used in SUMIF function or the SUMIFS functions. The INDIRECT function is not often very useful on its own (as you can simply point to the cell and find the number. Rather, the usefulness of the function is to use it together with other functions when the format of the variables is similar. To apply the SUMIF function using the example above, the following formula would be entered:

SUMIF(range of years in detail sheet, year criteria in annual sheet, INDIRECT(Range Name)

An example of using this technique to extract annual data from detailed periodic data is illustrated below:

Year	Show Comments	Balance	Row Num	Sheet	Cell Reference	2012	2013	2014
Capital Expenditure		FALSE	60	Financial Working Analysis	'Financial Working Analysis'!60:60	-	-	9,237.36
Operating Expenditure		FALSE	70	Financial Working Analysis	'Financial Working Analysis'!70:70	-	-	-
Revenues		FALSE	74	Financial Working Analysis	'Financial Working Analysis'!74:74	-	-	-
Capacity		TRUE	32	Financial Working Analysis	'Financial Working Analysis'!32:32	-	-	-
Generation		FALSE	46	Financial Working Analysis	'Financial Working Analysis'!46:46	-	-	-
Pre-tax CF		FALSE	76	Financial Working Analysis	'Financial Working Analysis'!76:76	- 162.50	- 1,950.00	- 7,739.49
IRR pre-tax						6.22%		
MIRR						6.78%		

Quickly and Elegantly Creating Graphs for Sensitivity Analysis

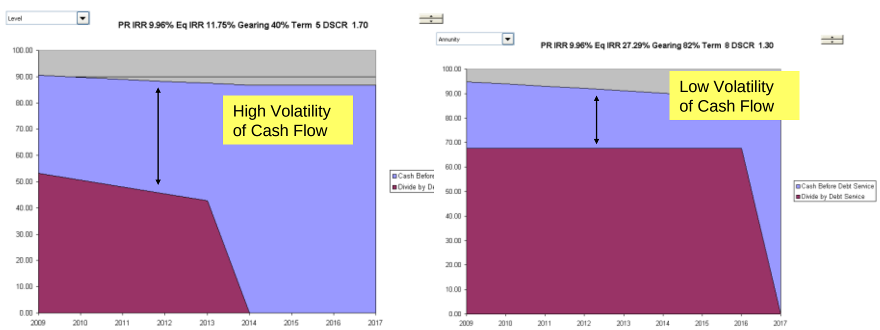
In creating a model, you can easily create thousands of graphs by selecting different variables and pressing the F11 key. Rather than presenting information that has marginal information for valuation, one of the arts in presenting valuation of equity and debt is to select graphs that effectively present risks and values in a summary page. For a project finance model, one of the essential graphs should be the cash flow available for debt service relative to the debt service over the operating life of the project. This graph not only shows risks to debt holders, but also illustrates the equity returns that are driven by the various structuring aspects of the debt including the DSCR, the debt tenor, the type of funding, the DSRA and so forth. For a corporate model, presenting historic and projected return on invested capital that isolates return from operations without financial effects such as the drag on earnings from surplus cash along with the growth rate effectively summarizes many of the important assumptions and results of the model. Both of these graphs could be obtained from the annual analysis discussed above and should be flexible with respect to different key dates in the model such as the terminal date or the

project life. It is more important to make these focused graphs flexible and effective than to present a hundred graphs that do not have much meaning in terms of risk and value.

The graph below illustrates presentation of cash flow and debt service that is essential a project finance model. Using the sensitivity graph one can see how low selected variables can fall in order for cash flows in a downside case to cover debt. If cash flows cannot cover debt service in a downside case (perhaps with some margin) then the amount and the terms of the debt should change. In order to quickly and effectively present and graph important cash flow and valuation statistics in the context of economic and financial structuring variables from the annual sheet or from other sheets, a few practical ideas are presented in the step by step analysis below.

Debt Capacity from Cash Flows with Different Volatility

- High Risk Cash Flows
- Low Risk Cash Flows



High Risk Project has higher margin, shorter-term and declining debt service. Low risk has flat debt service, and longer-term and higher IRR on Equity

Step 1: Use the F11 or ALT, F1 short-cut keys to make a quick graph

Select a variable from the financial model that presents cash flow or balances of a financial variable for presentation. Examples may include the cash flow and debt service as illustrated above, the return on invested capital, the balance of subordinated debt, or earnings per share. Either compute these variables in the annual sheet or transfer the variables from the detailed sheet as described above.

Collect the output variables in a separate section of the annual sheet in the model (you could name this section of the model something like “graph data.”) If you structure the rows and columns in this part of the model with the name of the output variable in the cell that is immediately to the left of the data and the x-axis of the graph (e.g. the year) immediately above the data, then you can press the F11 key and very quickly make a graph of the data (the F11 key should quickly become one of your favourite short-cut keys.) In excel 2007 and 2010 you can also use the ALT and F1 key to place the graph in same spreadsheet.

While the F11 key is a very nice simple method there are a few points that can make the process of making a graph much more flexible and effective:

- Create a variable that represents the x-axis above the data if the data is arranged in rows (and to the left of the data if the data is arranged in columns). For example, if you want years to be shown on the x-axis of the graph, then a row with the years should be above the rows of data to be graphed;

- Delete the name of the x-axis from the spreadsheet. This means that if the years are going to appear on the x-axis, then make sure the word “YEAR” is not to the left of the data. This is an important point that can make all sorts of presentations much easier. It is also essential when making an x-y chart;

- If there are blank columns between the data that is to be graphed and the titles of the data, then begin the x-axis data only in the year that is to be graphed. For example, if the graph begins in the year 2014 which is in column G and the titles are in column D, then enter the year above the column G and leave the other columns (E and F) blank; and finally,

- Shade the area of the data including the x-axis and the blank cell that is to the left of the data for the x-axis. After the data is highlighted press the F11 key (or the ALT and F1 key).

An example of setting-up data to effectively graph data using the F11 key or the ALT and F1 key is illustrated in the table below.

Note there is no title on the x-axis

Note there is no x-axis on the blank column

		2015	2016	2017	2018	2019	2020	2021	2022	2023	2024	2025	2026	2027
CFADS		15,916.48	21,329.84	20,311.36	20,149.40	20,246.75	20,357.65	20,441.11	19,987.40	19,877.77	20,065.56	20,082.38	20,124.85	19,913.18
Debt Service		12,243.45	16,407.57	15,624.13	15,499.54	15,574.42	15,659.73	15,723.93	15,374.92	15,290.59	15,435.04	15,447.98	15,480.66	15,317.83

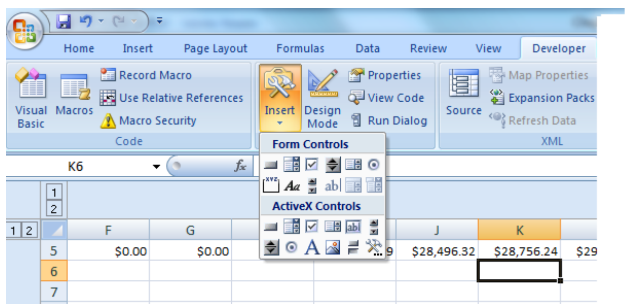
Shade area in the box and press F11 or ALT and F1

Step 2: Create a spinner button or other form to illustrate effect of selected variables

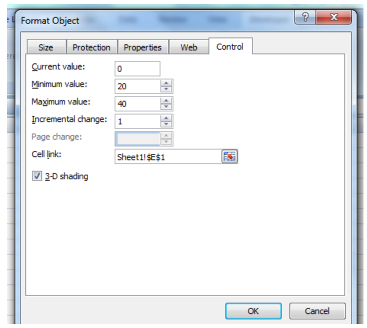
The idea of presenting sensitivity analysis is to show how an input variable affects the output variables that are chosen for presentation with something like a spinner button. The advantage of a spinner button or other form such as a combo box or a scroll bar is that you can keep the input variables in the input sheet, but you can also manipulate those input variables from other places in a workbook such as the summary page or a graph. For example, the input for operating expenses is still on the input page, but you if you use a spinner button then you can change the value of the operating expenses from either the summary page or from a graph or both. The disadvantage of using spinner boxes is that you can forget the original data that was used in the base case (resolving this problem is discussed in the scenario below).

To make a graph where a spinner button adjusts one of the input variables first choose one or more input variables that have an important effect on the rate of return, value or credit quality such as the oil price in the example above. To allow the variables to be effectively presented in the graph, you can use the spinner, scroll bar, check box or combo box forms available in excel. Unfortunately, there are a few quirks in excel that complicate this process, particularly when the graph is on a different page from the input data. A process for making the use of excel forms such as the spinner box more manageable is the following:

- In order to create a form such as a spinner box, use the view, toolbars, forms menu in excel 2003. In excel 2007 or 2010 you must include the developer tab on your ribbon. Once the forms are shown on the menu or the developer ribbon is on the screen, or you can use the insert controls tab from the developer menu in excel 2007 (make sure to use the control forms and not the activex controls) and insert the control on the sheet. Click on one of the forms (at the top) and then paint the form in the spreadsheet (this may be a bit more difficult for people born before 1965). The diagram below illustrates how to find Form Controls in excel 2007 or 2010.



- Once you have painted the form on the sheet, select the form with your mouse and right click on the form. Then select the FORMAT CONTROL option to attach the control to the input cell. The problem here is that the form (e.g. the spinner button) attaches to a cell in the particular spreadsheet where the form is placed. If you put the spinner box in sheet1, then a particular cell – e.g. E1 -- will be changed in sheet 1, but if you copy the spinner form to sheet 2, it will change E1 in the sheet 2 rather than sheet 1. This all but defeats the whole purpose of using forms like the spinner box because the idea of spinner button is to be able to change data in the input sheet without moving inputs to another sheet such as a summary sheet. It also means that you cannot copy the form to a graph if the graph is in a separate sheet. The solution to this problem of moving the spinner button and other form controls is to include the name of the sheet in the spinner box. This can be accomplished with a technique some have called the wind shield wiper method where you use the mouse to click on a different sheet from the sheet containing the cell link and then move back to the original sheet like a wind shield wiper. To use the windshield wiper approach, first click on different sheet (other than a chart sheet) and then move back to the current sheet so that the sheet name is included in the cell link is illustrated on the diagram below. Once you have created a form and attached it to an input cell in this manner, you can copy the form to other sheets and place the spinner box or other form directly in the graph.



Step 3: Add Titles to Graph with Summary Statistics

It is often useful to present a title on the graph that contains both text and statistics that represent important valuation outputs such as the internal rate of return on the investment, the minimum debt service coverage ratio, or the net present value of cash flows. In addition, you may want to show some of the input variables that are being changed with the control forms (as shown in the graph above.) As with the above process above, there are a few quirks in excel that complicate the process:

- When combining text and numbers on a graph, you must first enter all of the information that will be included in the title into a single row (or column) somewhere in your workbook (you can change the format of the numbers and make the title as long as you would like.) For example if you want to show the DCF value and the WACC, you may enter the following information on a single row:

  Enterprise Value: 4,556 Euro using WACC of 6.4%

- Once the information that will become the title has been entered in a single row or column, select the existing title on the graph (if there is not an existing title then you should create one.) Then press the equal sign from the keyboard and shade the single row or column where you entered the data. After shading the area in the non-graph sheet, press the enter key and the title should appear on the graph.

- If you would like the title on the graph to have multiple rows as in the oil price example above, you cannot press the enter key on the graph title to obtain a new line. Instead, you can create a new line character in the graph title where you entered the single line to be graphed. To insert a new line character, use the ALT and ENTER keys together at various cutoff points in the single row that you use for the graph title.

Optional Step 4: Make Dynamic Graph with the OFFSET Function

Sometimes you would like to make a flexible graph in which the length of the data in the graph can vary. For example, in a project finance model, the length of the concession agreement may vary; in a corporate model the explicit period may be flexible; and in an acquisition model the length of the holding period may change. To make a dynamic graph where the x-axis changes as a function of the model timing inputs, an effective technique is to use a flexible range name with the OFFSET function and then use the flexible range name as the source of data in the graph. This involves creating a range name with a length or width that changes as when various timing variables in the model such as the explicit

period or the construction period are changed. To create such a dynamic range name and a flexible graph, the first step is to understand the OFFSET function which has the following form:

=OFFSET(Reference Cell, Rows Down, Columns Across, Length, Width)

The first three arguments in the OFFSET function are required arguments and the last two are optional, but essential in creating a flexible range name. The first argument in the OFFSET function is the cell reference that will be used as a starting point. The second argument is the number of rows to go down from the starting cell reference and the third function is the number of columns to the right of the cell reference. For example the function OFFSET(A1,1,1) would result in cell B2 – it would start with A! and then go down by one row and to the right one column. The number of rows and columns can be negative implying to go up rather than down and to the left rather than to the right.

The important part of the OFFSET function in creating dynamic range names is not the first two arguments discussed above that allow you to move around, but rather the final two arguments relating to the width and the height. These arguments are generally used in the context of another function or as an array function. An example of how the height and width arguments can be used is shown in the two examples below where you would like to make an operation on a flexible range that varies according to some other variable in the spreadsheet:

= SUM(OFFSET(A1,2,3,2,5))

= NPV(10%,OFFSET(Cash flow reference,0,historic period,1,explicit period)

- The first statement begins with the cell A1 and then moves down two rows to row 3 and moves across three columns to the right to column D. This implies the starting point for the sum calculation is D3. The second statement begins with the first column which is a historic period and then moves across to the end of the historic data.

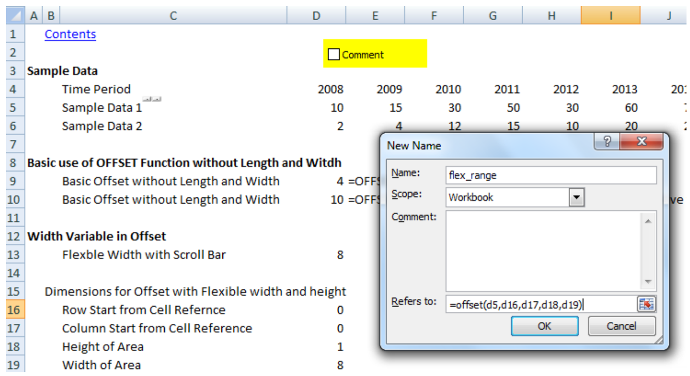
- Once the starting cell is established, the sum in the first statement covers two rows and 5 columns because of the last two arguments of the function. This means the area from D3 to H3 is summed – across 5 columns and it would also sum items in row 4. Therefore, the sum range would be D3 to H4. In the second statement, the NPV is computed for a single row that extends for the explicit period.

Once you understand the general way that the OFFSET function works, the process for creating a dynamic range name and chart involves the following steps:

- Step 1: Create a range name that will use the OFFSET function as an equation. To make a new range name you can press the CNTL and F3 keys (short-cut keys associated with range names use the F3 key).

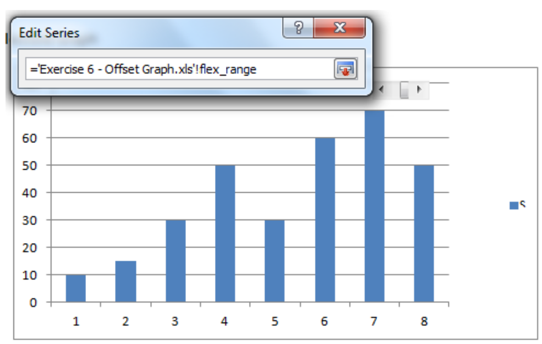
- Step 2: After defining the name of the range, select the range name and enter the OFFSET function in the "refers to" part of the range name menu. The key point is that you can use a function in a range name instead of simply referring to a cell or a range of cells. For example, if you want to make a range called flex_name that begins at D5 and has a width of 8, the offset function in the range name would not move away from D5 as the starting point. It would have a height of 1 and a width of 8 as illustrated below. (The parameters for OFFSET function – the row start, column start, the height and the width should generally be entered in as cells in the spreadsheet as shown in the diagram):

= OFFSET(D5,0,0,1,8)



- Step 3: Make a graph that will be soon be changed to include the flexible range name instead of a fixed area. To make the graphs you can use the F11 short-cut as described above.

- Step 4: After making the graph, edit it and select the data series option from the drop down menu. When editing the data series modify the range and replace the existing source with the range name but do not change the sheet name. This means that you should only replace the range after the exclamation point and leave in the name of the sheet before the explanation point. For example if A1:A5 will be replaced by the range name, do not change the sheet name to the left as illustrated below:



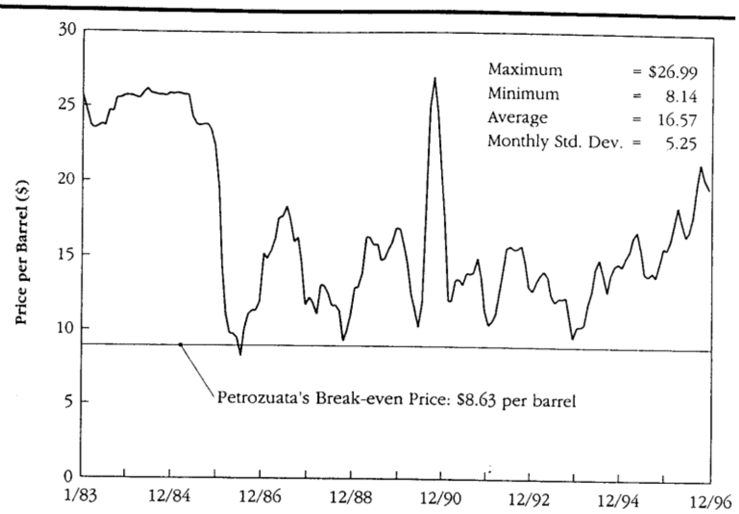
Before finishing the discussion of sensitivity graphs with spinner buttons and moving to break-even analysis, note that adding a whole lot of spinner boxes to inputs in a financial model can be a bad idea. The danger is that you mess around with a lot of different variables using spinner boxes and then you forget what the base case values were. For example say there is a spinner box for oil price variables, quantity variables, margin variables and other items. After playing around with graphs you will soon forget what the original variables were supposed to be. A solution to this problem is described below where the spinner boxes are combined with scenario analysis where a special custom scenario is

created that allows spinner boxes. If this technique is used, you can switch back to the base case or the downside case.

Chapter 17: Using Financial Models to Establish Break-Even Points for Key Input Variables to so that you do not need for Impossible Subjective Forecasts of Items such as Oil Prices

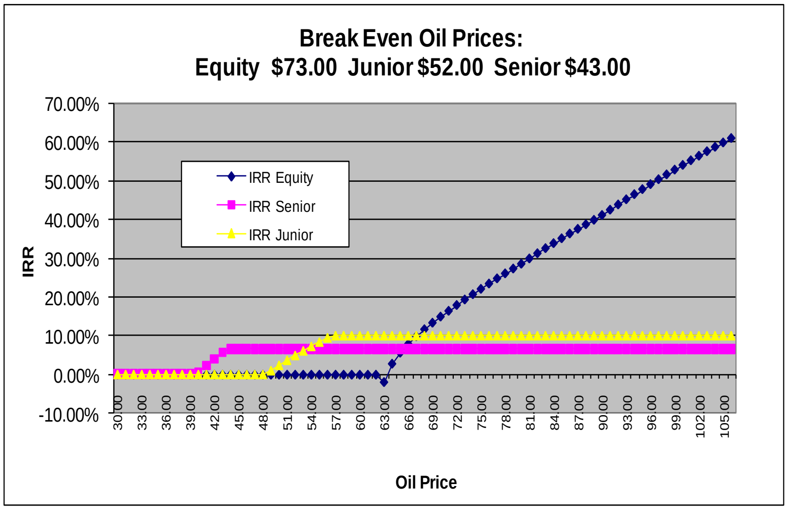
Value comes from cost structure, technological innovation, location, marketing strategy and other aspects of business strategy. In the long-run it does not generally arise from superior ability to forecast commodity prices, GDP growth, trends in interest rates or other economic variables. While these variables have important if not dramatic effects on value, one can argue that anybody who thinks he can forecast such variables is committing fraud. Instead of wasting time on consulting studies or marking reports that forecast these variables you can compute break-even values for the economic variables that are all but impossible to forecast. Break-even analysis is similar to sensitivity analysis, with a subtle but important difference that you can quantify risk as a single number that shows the cushion by which a variable can move before something bad or good happens. One of the supposed advantages of fancy statistical measures such as value at risk or the probability of default is that these numbers seemingly place all of the risk of an investment into a single number. By expressing risk as the break-even point a similar thing is accomplished, but in a more intuitive way. One of the famous cases of a valuation nightmare was the case of Eurotunnel where dramatic errors in the volume of traffic created massive losses on debt and equity capital. If you are an investor in one of the many tranches of junior debt to evaluate the risk that your debt cannot be paid you could construct a financial model of the project and then keep pushing the traffic volume down until it reaches the level at which your debt cannot be repaid and you lose money. Once you have computed this break-even traffic level, if the chance that the actual traffic level will be below the break-even traffic level is very low, you could be confident that your debt will be repaid. On the other hand if the cushion between the base case traffic level forecasted by the consultant and the break-even level is thin, then the likelihood of default on your debt may be quite high. The break-even traffic level has defined the risk of the project. Similar break-even measures could be developed for commodity prices, demand growth, cost structure and for different classes of investors – equity, senior debt, preferred stock and so forth. An example presented in a project finance context is illustrated below where the break-even oil price is compared to a level of oil prices that would produce a DSCR of at least 1.0.

FIGURE 1
MAYA CRUDE OIL PRICES



Source: Energy Information Administration.

The graph below presents break-even analysis using the oil transaction example above for senior debt, subordinated debt and equity. The three lines on the graph show the rate of return (the IRR) on senior debt, subordinated debt and equity while the x-axis lists a range in potential oil prices. The level at which the variables fall below the risk free interest rate is the defined as the break-even point as you could alternatively invest in risk free instruments. Different levels of break-even at which the senior debt, subordinated debt and the equity yield a return below the risk free rate is shown in the title of the graph using the method of putting the title on one line in somewhere in the spreadsheet described above. If the current oil price is $85/barrel, then the cushion for the equity break-even is quite thin, while the cushion for the senior debt is much higher because prices can fall all the way down to $43/barrel before the IRR on senior debt falls below the risk free rate. If the transaction used different levels of debt or equity, the break-even values would change; if more senior debt and less equity are issued to purchase the company, the break-even cushion would be thinner for the senior debt and higher for equity. Making a break-even presentation like this may not define the required return necessary to compensate for taking risk, but it does effectively display risks relative to returns for alternative investments and allows you to make a judgmental assessment of risk and return.



While general intuition behind a break-even analysis is simple, there are a couple of issues in computing and presenting sensitivity analysis that are a little tricky. The first issue is establishing an effective criterion for purposes of deriving the break-even point. Choosing the criteria at which the break-even is assumed to occur – when the DSCR equals one or the IRR equals the risk free rate interest rate seems pretty obvious. However the break-even criteria can require judgment and deep understanding of financial ratios. A second issue in working with break-even analysis involves developing mechanical techniques to automatically change the break-even level whenever a structural variable changes in a financial model. When you vary the level of senior and/or subordinated debt or the interest rate or the purchase price in a transaction, the break-even points will change and it is useful to work through the process of automatically changing the break-even points.

Difficult Questions in Establishing Break-even Criteria when Making Analysis with Financial Models

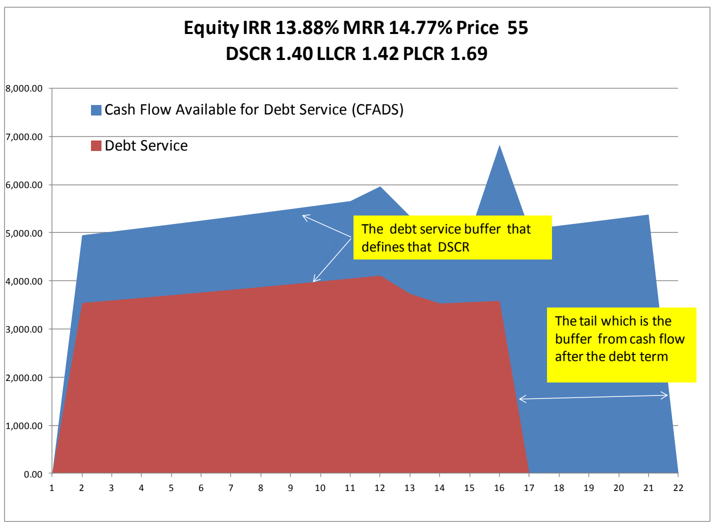
The process of establishing a criterion for determining the break-even point is not as simple as it seems at first blush. Consider the perspective of senior lenders in a transaction and assume the most important risk variable is the projected growth rate. In this case the break-even analysis is the lowest acceptable growth rate which is acceptable before something bad happens to the lenders. In a project finance analysis, one approach is to measure how low the growth rate can fall be before the minimum debt service coverage ratio becomes 1.0x – the level at which cash flow just covers debt service on a year by year basis. The 1.0x criterion measures at what level cash flow is insufficient to meet the required payment, which seems to be a natural basis upon which to compute the break-even point if project default is the issue. The minimum debt service coverage however may not represent the actual loss on debt as a company could be able to ultimately re-pay debt after missing a debt service payment because there may be enough cash flow to re-structure the debt and still be fully repaid. An alternative break-even analysis is computing how low the growth rate can go until debt cannot be repaid at the end of a project. A third way to look at the issue is to evaluate how low the variable can fall until the IRR on debt falls below the risk free rate. This approach uses the notion that an investor has an option to either invest in a risk free security or risky investments and the break-even should shown what risks are taken to achieve a higher return than the risk free rate. These three different criteria can lead to substantially

different break-even points. Similar issues arise with respect to break-even points for equity investors, for acquirers in merger analysis, for government agencies and other decision makers.

The graph below of cash flow available for debt service and debt service in a project finance transaction illustrates the difference between evaluating break-even using a debt service coverage ratio and break-even points derived from a criterion that accounts for the total cash flow over the lifetime of a project. In a project financing, much of the analysis from a debt perspective is about buffers or cushions in cash flow relative to debt service. A primary buffer is the area between the cash flow and the debt service over the debt tenor that defines the debt service coverage ratio (DSCR). The second buffer is the area of surplus cash flow after the debt has been repaid which is called the tail. If the break-even is derived from the DSCR, then the risk buffer from the tail is ignored. To account for the value of the tail, the project life coverage can be used. This ratio is defined as the present value of the cash flow using the interest rate on debt divided by the present value of the debt service. As the present value of debt service equals the value of the loan, the DSCR and the PLCR can be defined as follows:

DSCR = Cash Flow Available for Debt Service/Debt Service

PLCR = PV(Cash Flow Available for Debt Service)/Loan

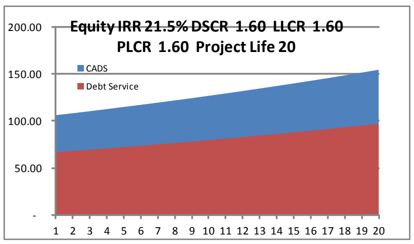


A third ratio can also be computed which evaluates the ability of cash flow to service the loan over the life of the loan called the loan life coverage ratio or the LLCR. This ratio ignores the value of the tail but evaluates the safety buffer of the loan over the entire period of the loan. To compute this ratio, the formula is the same as for the PLCR, except that the cash flow is computed over the life of the loan rather than over the life of the project.

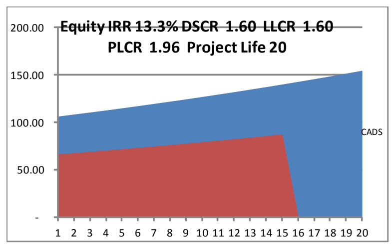
The three graphs below demonstrate the difference in computing break-even points using different financial ratio criteria. In the first case there is no tail and the debt service buffer (the DSCR) remains constant over the life of the project using a sculpting method of debt repayment. Here the DSCR, PLCR and LLCR all result in the same break-even points. In the second example, the debt service buffer is constant, but there is a tail. In this second case, the LLCR equals the DSCR but the PLCR is greater than the DSCR or the LLCR because of the risk buffer from the tail. In the third case, the debt service does not follow the cash flow and the buffer increases over the term of the loan. Here, the minimum

DSCR in the first year is less than the DSCR in later years. If one uses the minimum DSCR – the point at which a default occurs – to gauge the break-even point, the analysis will be different to using the LLCR or the PLCR. If the DSCR is used as the break-even point there is a single year default, but the loan can eventually be repaid. If the LLCR is used in the break-even analysis, the break-even is defined as the point where a single year default may occur, but the loan is repaid by its scheduled repayment date. If you would like to find out how low a variable can go before a loss occurs on the loan, the PLCR should be used as the break-even criteria.

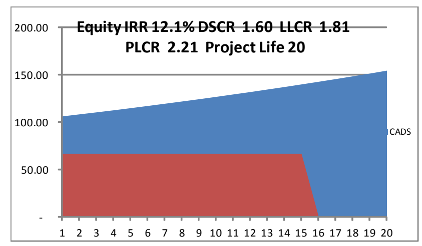
Case 1: Sculpting; No Tail



Case 2: Sculpting with Tail



Case 3: Annuity with Tail



Establishing break-even criteria from a lender perspective in corporate finance can be even more challenging. In corporate finance the lender must evaluate how bad things can get before re-financing becomes impossible. To do this some financial ratio such as the Debt to EBTIDA can be measured and then judgment must be considered as to how high the Debt to EBITDA can get before re-financing cannot occur.

Mechanics of Using Data Tables to Automatically Compute Break-Even Points

After establishing the financial criteria that will be used to compute the break-even point, the mechanics of computing and presenting the break-even points can be developed. Implementing break-even analysis can be simply thought of as a subset of sensitivity analysis discussed above. In the graph of the cash flow waterfall and oil price presented in the sensitivity analysis section, the break-even oil price can be determined by simply pushing down the oil price until the IRR on the subordinated debt falls below the risk free interest rate (if that is the break-even point that was chosen). Another method of computing the break-even point is simply to use the goal seek tool in excel to derive the required level of the input variable that just meets the break-even point. A third approach is to use the data table tool that lists increments of the break-even variable (e.g. the growth rate in demand) next to values produced of the criteria variables (e.g. the debt or equity IRR.) This method allows one to compute break-even points for different investments (and/or for different criteria) and present the break-even points in a summary output without re-running the goal seek process each time the structure of the different securities as shown above for senior debt, subordinated debt and equity. The step by step instructions below explain how to compute break-even values using the second technique using data tables.

Step 1: Create a one-way data table

The first step in the process of automatically presenting break-even points is to use the data table tool to create a one way data table. With a one way data table, the sensitivity values are increments of the break-even variable (such as oil prices) are listed and a set of values for the criteria variables are computed for each of the increments. The data table tool described in this section (accessed from the

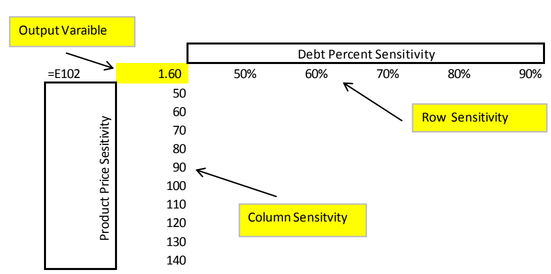
data menu in excel 2003 and from the what-if button in the data tab in excel 2007) is both one of the best and the worst features of excel. The good aspects of data tables are that you can quickly create scenarios and produce different break-even values for different transaction structures. The bad part is first that the data tables must be in the same sheet as the input variable – in this case the sheet that contains the oil price, the growth rate, the traffic demand or other break even variables. Another bad thing about data tables is that they can seriously slow down excel (unless the automatic except data table option is used.) An alternative to using data tables using some simple VBA code is explained below.

In explaining how to use data tables one should understand the difference between one-way data tables and two-way data tables. Data tables perform a sensitivity analysis where an output variable (such as DSCR, Debt/EBITDA, or the IRR) is presented next to a range of input variables. For example, the return on invested capital could be presented next to gross margin assumption in a corporate model. The difference between a two way table and a one way table is that:

- A two way can present only one output variable but allows two input variables to be presented in the sensitivity.
- A one way data table only allows one input variable to be used in the sensitivity analysis but can display the effect on as many output variables as you want.

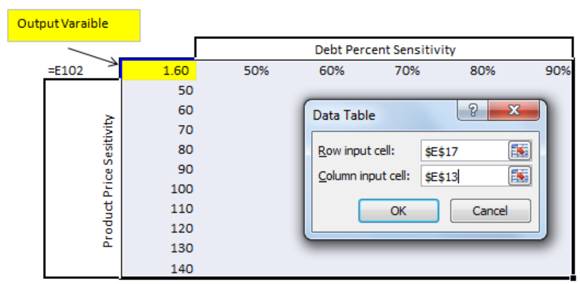
Creating a two way table is more intuitive than the one way data table and is explained first. One way data tables involve:

- Select two sensitivity variables somewhere in the same sheet as the inputs. You must enter the variables in a particular structured format where one of the sensitivity variables is typed in a row and the other is typed in a column with a space in the upper left cell as illustrated below. The blank cell in the upper left hand cell between the row and the column cell must have a formula (labelled as the output variable below) that depends on the two sensitivity variables. In the example below, the row sensitivity is the debt commitment, the column sensitivity is the product price and the output variable in the top left cell is the PLCR computed in the model.



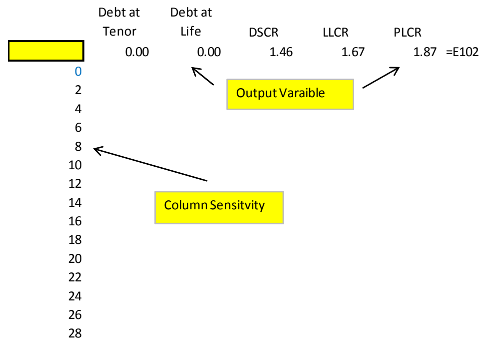
- After carefully typing the structure of the sensitivity variable as shown above with the output variable at the top left, you can fill in the sensitivity analysis using the data table tool. In the example above, the table is completed through changing input variables in the model from 50% to 90% for the debt percent input assumption and changing the product price input in the model assumptions from 50 to 140. The variables that are typed for the row and column sensitivity are not connected to anything in the financial model and do not yet have any effect on the output variable. To complete the data table, the excel program must know

how to connect the sensitivity variables listed to the left and above the data table to the input assumptions. This is done by first shading the area of the table which includes the output variable, the sensitivities and the results area (it must not include the titles to the left of the column sensitivity or on top of the row sensitivity). In the example above, begin shading yellow cell and shade up to the 90% and down to the 140 and do not shade the titles of the sensitivity variables. After shading the area, select the data and table from the excel menu (using the what-if analysis in excel 2007 or 2010). Then, you will be asked to type in a row input and a column input. This is where you connect the sensitivity ranges in the row and the column to inputs in the model. You must find the single variable in the model inputs that corresponds to the sensitivity row and then find the single variable in the model inputs that corresponds to the sensitivity column. At first this may seem somewhat confusing, but you will get used to it after trying it a few times if you remember that the data you entered in the rows and columns has no connection to the model. The process of inputting row and column inputs is illustrated below.



For developing a break-even analysis, a one-way data table is more useful as you are generally concentrating on one input rather than two input variables. The reason this type of table is called a one way table is that one input variable is analysed rather than two. Further, as illustrated in the example above where the break-even for senior debt, subordinated debt and equity are presented, you may want to evaluate more than one output variable. Unlike a two-way table, a one way table can be used to create multiple output variables because of the two-dimensional structure of a spreadsheet. The process of creating a one way table differs from a two way table as follows:

- Unlike the two way data table described above, you must structure the one-way table through arranging the output variables and the single input variable in a different way. With a one-way table you leave the square above the column sensitivity and to the left of the output variables blank (shown in yellow cell on the diagram below). The one or many output variables are then entered to the right and above the single column through linking to the financial model outputs. When you are finished structuring the table, there should be a blank cell to above and to the left of the sensitivity variables and the output variables. An example of the structure of a one way data table is illustrated below.



After carefully setting up the one-way data table as illustrated above, you shade the data table as with the two way table and include the blank cell sell illustrated in yellow above in the shaded area (again, you must not include the titles of the row and the column in the shading). After shading the area, go to the data table menu and you are asked again to enter a row and a column input cell. The non-intuitive aspect of a one way data table is that you should leave the row input blank because the items across the row are outputs and not inputs. For the column input cell, you find an input assumption in the financial model. You can use the same process by entering the output variables on a column and the sensitivity variables and then you must enter something in the row input and nothing in the column input. After finishing entering the data table, it should look something like the diagram at the end of the next paragraph.

Step 2: Use MATCH and INDEX functions to find beak-even points

Once the data table is created that lists the input sensitivity variable (such as oil price) next to the output variable(s) (such as DSCR, LLCR and PLCR), the remaining task is to look into the data table and find the break-even points for the alternative variables. To find the break-even value for each criterion, two excel functions which are like brothers – the MATCH function and the INDEX function -- can be used together. (The LOOKUP function could also be used.) The INDEX function has already been introduced where you use it to pick out an item from an area given a row and a column number – INDEX(area, row, and column). The MATCH function is useful together with the INDEX variable because you if produces a row number or column number once you enter a criteria.

To use the MATCH and INDEX function in finding the break-even point by first use the MATCH function to find the row number of the table that corresponds to the break-even criteria (in the example below the criteria is 6 %.) The MATCH function would be written as follows for the table below:

Row of Break-even = MATCH(Break Even Criteria (6%), column of numbers in data table)

Once the row number has been found with the MATCH function, the INDEX function can be used to find the break-even value of the input variable that is associated with the row number. To do this, you first shade the input sensitivity value of the data table (in the example below, the oil price) and then you enter the break-even row from the match function as follows:

Break-even value = INDEX(sensitivity row range, break-even row from MATCH)

This is an input variable for the oil price that must be in this sheet

Input Variable - Oil Price 82

Set up the data table exactly like this with the formulas one row up and one to the right

		IRR Equity	IRR Senior	IRR Junior
	Hurdle	6%	6%	6%
	Row Num	26	15	19
	Break Even	65.00	43.00	51.00
		31.4%	6.77%	10.27%
OIL PRICE SENSITIVITY	35.00	N/M	0.00%	0.00%
	37.00	N/M	0.00%	0.00%
	39.00	N/M	0.00%	0.00%
	41.00	N/M	2.40%	0.00%
	43.00	N/M	5.75%	0.00%
	45.00	N/M	6.77%	0.00%
	47.00	N/M	6.77%	0.00%
	49.00	N/M	6.77%	1.00%
	51.00	N/M	6.77%	3.76%
	53.00	N/M	6.77%	6.25%
	55.00	N/M	6.77%	8.58%
	57.00	N/M	6.77%	10.27%
	59.00	N/M	6.77%	10.27%
	61.00	N/M	6.77%	10.27%
	63.00	-1.91%	6.77%	10.27%
	65.00	5.69%	6.77%	10.27%
	67.00	9.95%	6.77%	10.27%
	69.00	13.44%	6.77%	10.27%
	71.00	16.51%	6.77%	10.27%
	73.00	19.44%	6.77%	10.27%

Step 3: Attach Spinner Boxes for Sensitivity Analysis

When the data table is created in this manner, the value resulting from the INDEX function can be presented in the summary sheet or the dashboard. When you change a structuring variable such as the level of debt, the data table changes and the MATCH and INDEX values also change (the formulas must be set to automatic or a simple macro should be recorded from pressing the SHIFT and F9 keys. To create a simple macro, simply begin recording a new macro and then press the SHIFT and F9 keys. After the macro is created, attach a spinner box to one of the structuring variables such as the debt to capital in a leveraged buyout transaction and then add the macro to the spinner button by right clicking on the button.

Creating Data Tables using VBA Rather than using the Data Table Tool

Data tables can be useful tools in break-even and sensitivity analysis. However despite the benefits of the data table tool, it can present a number of difficulties including: (1) the data table cannot be presented in a sheet other than where the inputs are used in the data table; (2) a data table cannot be used together with a goal seek or macros in the model; and (3) a data table cannot be used from another table where the inputs to a second data table come from the first data table. To address these issues a macro can be created in lieu of the data table. The key to creating a data table with VBA is to use a FOR NEXT loop together with the CELLS function. If you get used to these options in writing VBA code, many things become possible and easy. To describe how to combine a FOR NEXT loop with the CELLS function, a simple one way data table is introduced with only one output variable. Then, additional complexities are added through creating a one way table with multiple output variables and finally, a two way data table.

To create a one way data table with VBA with one output variable, you need one FOR NEXT loop that repeats calculations over the number of rows in the data table. You also should use the CELLS function in VBA that allows you to define either an output cell or an input cell given a row and a column, In writing the code, you change the input variable in the financial model for each row increment using the CELLS function with the row number defined in the FOR NEXT loop. Finally, you list the output variable for each row increment again with the CELLS function. When entering data to create the table, you only

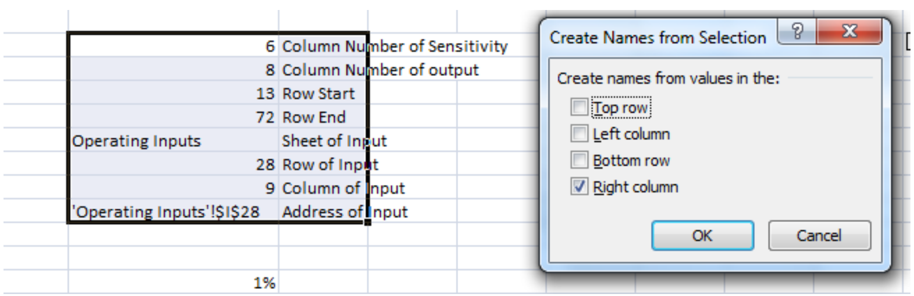
need to make a list of the input sensitivity for each of the 20 rows. The FOR NEXT loop would have the following form where the variable row could have any valid name.

```
For row = 1 to number of rows
        Range(financial model input) = CELLS(row, column number of sensitivity variable)
        CELLS(row, column number of break-even output) = Range(financial model output)
Next row
```

To make the data table flexible, you can enter range names for the column number of the sensitivity variable, the column number of the output variable and the number of rows. This can be accomplished by using the COLUMN and COUNT functions along with SHIFT, CNTL, F3 for naming ranges as illustrated in the equations below:

Row Start = ROW(first row of sensitivity)
Row End = ROW(last row of sensitivity)
Column number of sensitivity variable = COLUMN(One of the sensitivity variables)
Column of sensitivity variable output = COLUMN(column where you want output)

The excerpt below includes a cell reference for the input using the ROW, COLUMN and ADDRESS functions as well as a created sheet name function (using cell_reference.parent.name). The names of the cells are placed to the left of the data. After typing the names, the SHIFT, CNTL, F3 short cut is used to create the names.



The code below illustrates how a flexible data table can be created. The initial value is retained at the beginning of the program and replaced at the end of the program. Each item in the program comes from a cell reference in the sheet so that rows and columns can be inserted and deleted. The status bar is adjusted to display the row that is being computed.

```
Sub OneWayTable()

start_value = Range(Range("Address_of_Input"))

For Row = Range("Row_Start") To Range("Row_End")
   Range(Range("Address_of_Input")) = Cells(Row, Range("Column_Number_of_Sensitivity")) * 1000
   Cells(Row, Range("Column_number_of_output")) = Range("output")
   Application.StatusBar = " Row " & Row
Next Row

Range(Range("Address_of_Input")) = start_value
Application.StatusBar = False

End Sub
```

To create a one way data table with more than one output variables, you can simply add additional output variables and define more column numbers where the output variables will be placed. If you want to make a two way table, the trick is to use two FOR NEXT loops: the first loops around rows as above and the second FOR NEXT loop works through the columns. The number of rows is the length of the column sensitivity and the number of column is the length of the row sensitivity. The code below demonstrates the technique of using two loops. You can add range names and make the two way table flexible as in the case above.

```
For row = 1 to number of rows
        Range(column input) = cells(row, column of sensitivity variable)
        For column = 1 to number of columns
                Range(row input) = cells(column, row of sensitivity variable)
                Cells(row, column) = range(output)
        Next column
Next Row
```

Although the break-even analysis presents risk as a single number, there are a couple of problems with the technique. First, the break-even analysis does not measure what happens when a cohesive set of variables that can move together. A rather obvious example is that in an extended recession, sales and prices may decline in tandem. All of the perfect storms that seem to occur much more than would be expected if variables would really move independently demonstrate this point. Second, the analysis depends on a single value of the break-even variable – for example, the oil price is assumed to stay at the constant level over the term of the modelling period. This means that use of a break-even analysis is difficult to accomplish when a variable is volatile and the price can have sharp upward and downward moves.

Chapter 18: Constructing Flexible Scenario Analysis that Measures Exposure of Equity and Debt Investors to Carefully Constructed Downside and Upside Outlooks under Different Transaction Structures

An important alternative to using beta or other statistics to measure risk is to ask a banker how much he will lend to a company. The debt capacity not only measures credit risk but as the amount of debt defines the variation in equity cash flows, it also is central to equity and asset valuation. Simply put, if a banker with no vested interest puts his stamp of approval on a project, the investment will probably be made. If the banker concludes that the project is too risky, an equity investor has a very strong signal that something is wrong (so much for Modigliani and Miller). In determining how much debt can be supported by an investment, there is no single statistical formula as no one really knows how banks and other lenders make decisions. However one oft-mentioned reasonable theory is that bankers come up with a carefully considered downside case and then make sure there is some remaining buffer of cash flow relative to debt service in this downside scenario. To illustrate the notion of using a downside case to develop debt capacity, consider the example of a wind farm in Germany that receives a fixed "feed-in" tariff from the state and can hedge operations and maintenance risk using contracts. With energy prices and expenses hedged, the primary remaining risk is the possibility of electricity output from wind being lower than expected. When addressing this production risk, consultants construct probabilistic forecasts that (supposedly) address wind speed risk, errors in a host of statistical parameters that measure relationships between different wind speeds and a variety of technical parameters associated with the wind turbines. To determine how much debt a project such as this can support, a one-year P90

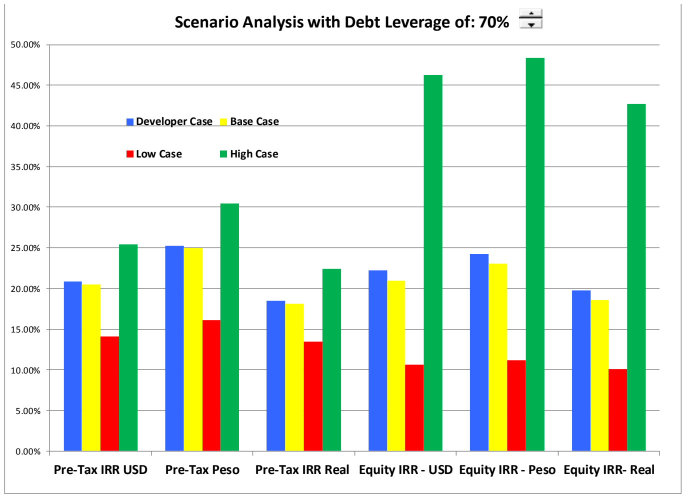
production level for wind speed could be as a downside case meaning that in any one year there is a 90% chance that the wind will be higher than this level. Then, to compute the amount of debt, the bank may add a 15% buffer on top of the expected cash flow in this downside case implying a DSCR of 1.15x. This example demonstrates how scenario analysis – in particular the downside case -- drives the debt sizing and hence valuation decisions in real world risk assessment. The size of the debt and the viability of the investment depends not on the base case (which may be the P50 case), but the P90 downside case assumptions.

The difference between scenario analysis and the sensitivity and break-even techniques described above is that in a scenario analysis, multiple different input variables are changed at the same time rather than focusing on a single variable. In developing a scenario analysis, one must come up with a comprehensive outlook with respect to a number of economic variables rather than simply seeing how much one variable can change before something bad happens. If a company can easily survive a downside case – it can repay debt out of cash flow or easily re-finance debt maturities -- without cutting dividends and without changing capital expenditure policy, then the credit rating should be above investment grade. If the company cannot survive the downside case – perhaps meaning it cannot re-pay or re-finance maturing debt without dramatic changes in capital expenditures and other items -- then the credit rating should be below investment grade (BBB- in the Standard and Poor's scale.)[2] In a project finance context, the downside case can be used to establish covenants, debt service reserves and other elements of a transaction as shown in the table below:

Scenario	**Use in Project Fiance Model**
Base Case	Establish the schedule for repayment of debt
Downside Case	Establish the level of debt service reserve Establish the covenant levels Develop repayment flexibility
Upside Case	Establish cash sweep mechanics Develop pre-payment structure

An illustration of creating a scenario analysis is shown on the graph below for a corporate valuation. The base case, downside case and upside case are presented using four different terminal valuation techniques in order to illustrate the potential variation in the stock value. A well structured scenario analysis could present the variation in any other variable in the financial model such as the DSCR, LLCR, PLCR, or the IRR in a project finance model. Further, in presenting the analysis, structuring variables or timing variables should be displayed along with the scenarios. In the case of shown below, the value is presented using different explicit periods and different WACC assumptions (with the spinner box). In a project finance model, the scenarios may be presented with different assumptions for the debt commitment, debt tenor and other debt structuring variables.

2 Fundamentals of Credit Analysis, McGraw Hill.



The most important issue in developing scenario analysis is determining assumptions for the level of various variables in the alternative cases. It is difficult enough to come up with base case assumptions which are intended to the expected or most likely outcome (not the budgeted case presented by company management which often contains an optimism bias related to completion of projects and implementation of business strategy.) In developing a downside case, one must not only come up with reasonable values for the variables, but also some sense of what the downside case is supposed to represent. For example, banks sometimes define the downside case as a 20% probability scenario, but nobody really has any idea of whether the set of variables really has a 20% chance of occurring. It is very easy to come up with a pessimistic set of variables; it is much more difficult to develop a set of consistent variables that are quite negative and have a relatively small chance of occurring.

In the wind project example discussed above, the downside is relatively easy to develop if you believe consultant forecasts because the risks of different wind conditions and technical functioning of a turbine can be observed from objective past data which should not deviate from historic data, meaning that future possibilities for wind speeds can be taken from a given distribution. On the other hand, construction of a downside case for the sub-prime mortgages would have been much more difficult to establish. To come up with this scenario one would have to project future housing prices, future income levels, the relation between housing foreclosure and income and many other economic variables. Unlike projection of the wind speed in the case of the German wind turbine, historical observations of past data may not be very useful in forecasting what could happen in the future. The difficulty in coming up with this sort of judgmental analysis required for scenario analysis is probably one motivation for developing mathematical representations of variables which supposedly measure the probability of the scenarios.

Fitch rating agency publishes ranges in values that it uses for a base case and downside scenario. For example in making a downside analysis of a conventional electricity thermal plant with a fixed price contract (a PPA) the following assumptions are used:

Indicative Rating Case Stress Levels for Thermal Projects

	Change from base case		
Project stresses (%)	**Coal plants**	**Combined cycle**	**Peaking plant**
Heat rate	+1 to +5	+1 to +5	+1 to +5
Availability	-1 to -10	-1 to -10	-1 to -10
O&M costs	+5 to +15	+5 to +15	+5 to +15

Source: Fitch

Mechanics of Scenario Analysis

The remainder of this section addresses mechanics of adding a scenario analysis to a financial model in which different assumed cases can be easily added to a model and the scenario analysis can be tested and presented along with alternative structuring variables. Mechanical construction of scenario analysis is certainly not as important as coming up with the appropriate values to use in the different cases – in particular the downside case. But creating a well structured and effectively presented scenario analysis can allow you to spend more time on the more important issue of debating a downside case. The type of scenario analysis recommended in this section is a technique where one can adjust structural variables such as the size of debt and view what happens in alternative base, low and high cases. In describing the scenario analysis a simple method is described on a step by step basis where the scenario analysis should also be structured so that you can easily add a master scenario page to a model and choose a host of different variables including variables that have different values over time. Finally, the scenario analysis should allow you to use spinner buttons and other sensitivity features without losing track of the values that were established in the base case. A well designed scenario analysis should include the following:

- The ability to add multiple different scenarios and evaluate multiple different output variables;
- The potential to change transaction structuring variables and evaluate what happens across multiple scenarios;
- The addition of a separate master scenario page that can be easily be adjusted to change input variables, output variables and scenarios;.
- The potential to include variables that are structured as with changing values over time (time series variables) such as changing inflation rates, growth rates, interest rates and other items;
- The flexibility to include a special case where one can use spinner buttons and drop down boxes for sensitivity analysis and at the same time re-set the variables to the base case and other cases, and:
- The capability to extend the scenario analysis into a tornado diagram and a spider diagram (described in the sections below).

All of these mechanical capabilities can be accomplished by using the INDEX function in excel together with the DATA TABLE tool (or alternatively, a small VBA program as described above.) One can use the scenario manager in excel to make scenario analyses which allows you to change a number of different variables and create a report containing multiple output variables. Problems with the scenario manager approach include: (1) new scenario pages must be created each time a change in the structure of the transaction is made and one must re-run the scenario manager each time the scenario manager is run; (2) the titles of the input variables and output variables must be manually adjusted; (3)

the input variables must be entered into a data form rather than on a spreadsheet; (4) the scenario manager is difficult to manage when input variables are located on different pages; and (5) one cannot make drop down boxes to manage and graph the scenarios. Given these problems, using the INDEX function and the data tables is described below on a step by step basis.

To explain the mechanics of developing a scenario analysis that uses the INDEX and the DATA TABLE technique, a basic example using variables that do not vary over time; without the flexibility to add spinner buttons and without addition of tornado diagrams or spider graphs is presented first. Once the fundamental approach is established, additions of the items that are slightly more complex are added. At the end of the section, a technique using VBA instead of DATA TABLES and linking variables is described.

Step 1: Set-up a Master Scenario Page with a Scenario Number

To create a scenario analysis, you can first add a new page and then enter a scenario number somewhere at the top of the page. A basic concept behind creating any scenario analysis is that various scenarios should be assigned a number and that a scenario number should be input to define which scenario is being implemented. The scenario number will be the row or column number for implementing the INDEX function; it will be the cell link for the drop down or combo box that allows different scenarios to be operated from alternative places in the model; and, the scenario number is the becomes the column or row input variable for the data table. After the scenario number is input the scenario analysis works by using the scenario number to identify variables that will be used running the model (the scenario variables will replace inputs in the model). If scenario number one is used, then variables defined on the first row or column of the scenario inputs as illustrated on the table below where the scenario number is placed at the top of the page and the first row of data is also presented at the bottom of the table. The notion of using a scenario number is similar to the manner in which time series variables with a changing structure over time were discussed in Part 1.

Step 2: Enter data for different input variables in a scenario format.

The most tedious part of making a scenario analysis is entering data for different cases. This involves entering the base case, the downside case, the upside case and a set of different sensitivity scenarios in a different rows with values for different assumptions (e.g. price, demand, cost structure) in separate columns. An example of entering scenario inputs is shown in the figure below that illustrates a completed scenario analysis (the example uses columns for the scenarios and rows for the input variables.) There is no limit to the number of scenarios that can be added in this process – the example shown in the table below includes a number of scenarios in which only one variable is changed from the base case.

Scenario Number	1		Developer Case						
Variable Number	1	2		3	4	5	6	7	8	9
	Construction Period	Useful Life of Plant	PPA Life	Average Plant Production MW per Hour	Average Hours of Production per Day	Availablilty Factor	General U.S.Inflation Rate Code	General Mexican Inflation Rate Code	Real Change USD/Mexican Exchange Rate	Base Construction Cost of Equipment
Developer Case	15	40	20	35	24	86.50%	1	1	1	44,400,000
Base Case	15	40	20	35	22	86.50%	1	1	1	44,400,000
Low Case	20	30	20	33	20	86.50%	2	2	2	50,000,000
High Case	15	40	20	35	24	86.50%	3	3	3	40,000,000
Construction Period	20	40	20	35	24	86.50%	1	1	1	44,400,000
Useful Life of Plant	15	30	20	35	24	86.50%	1	1	1	44,400,000
Average Plant Production per Hour	15	40	20	33	24	86.50%	1	1	1	44,400,000
Average Hours of Production per Day	15	40	20	35	20	86.50%	1	1	1	44,400,000
Availablilty Factor	15	40	20	35	24	83.00%	1	1	1	44,400,000
General U.S.Inflation Rate Code	15	40	20	35	24	86.50%	2	1	1	44,400,000
General Mexican Inflation Rate Code	15	40	20	35	24	86.50%	1	2	1	44,400,000
Real Change USD/Mexican Exchange Rate	15	40	20	35	24	86.50%	1	1	2	44,400,000
Base Construction Cost of Equipment	15	40	20	35	24	86.50%	1	1	1	50,000,000
Contingency	15	40	20	35	24	86.50%	1	1	1	44,400,000
Consumables per kWh	15	40	20	35	24	86.50%	1	1	1	44,400,000
Custom Case	15	25	20	35	24	86.50%	USD Inflation	Mexico Inflation	1	44,400,000
Developer Case	**15**	**40**	**20**	**35**	**24**	**87%**	**1.00**	**1.00**	**1.00**	**44,400,000**

Step 3: Use INDEX function to find the input data associated with that scenario number

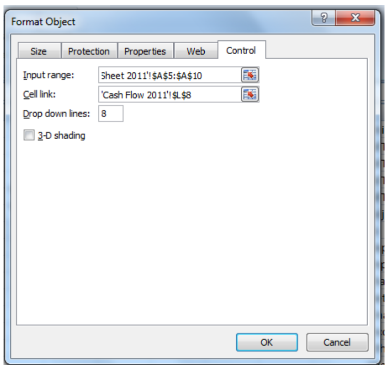
The central idea behind the scenario approach is to use the scenario number to change a comprehensive set of inputs in a model. The excel function that can accomplish this is the INDEX function. If the data is entered across a column with different scenarios listed in rows as shown in the above example, then the array entered in the INDEX function is the list of rows and the row number that selects the value is the scenario number. The result of the INDEX function is shown on row shaded in yellow at the bottom of the table. The values are the same as the first row of the table because scenario number 1 is selected. In creating the index function, the scenario number should be fixed with the F4 key as illustrated below.

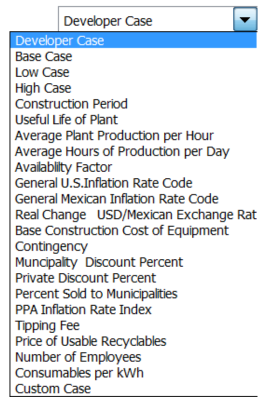
INDEX(column range, scenario number fixed with F4)

Step 4: Add a drop down box to enable running scenarios from different places in the workbook

To present different scenarios on graphs, on the summary page or in the financial model itself it is often convenient to include a dropdown box that lists the alternative scenarios. A dropdown box is illustrated next to the scenario number on the above diagram. Adding a drop down box is similar to the process of adding a spinner box described above for making flexible sensitivity graphs. In the case of the dropdown box, you select the developer menu, go to the insert option in the middle of the page and paint a combo box somewhere on the scenario spreadsheet page. To create a dropdown list that contains the scenario names, right click on the combo box and the select the FORMAT CONTROL option as was the case with the spinner box. The combo box requires and input range which is the list of items that will appear in the dropdown box. In this case the dropdown list should be the names of the various scenarios. As was the case with the spinner box, the “windscreen” method of moving to another sheet and then back to the scenario sheet should be used to assure that the sheet name is included in the input range. The diagram below illustrates the process of entering the input range and the cell link on the left hand side and the results of applying the dropdown box to the scenario list. The input range only works if the labels are listed in a single column across different rows (it does not work for a multiple columns listed in a single row). After selecting the input range, a cell link is entered that simply lists the result of the drop down selection in terms of the number of the row. In this case, the cell link should be the scenario number as this number also corresponds to the scenario list. Almost every time you

present a dropdown box you will probably use the INDEX function because once you have the cell link which is a counter for the rows, you can use this number to find elements associated with the list from the INDEX function. The cell link is simply an input as was the cell link in the spinner box. Note that by using a dropdown box you can still directly input the scenario number manually in the cell (which changes the selection of the combo box.) Alternatively you can select and item in the combo box which changes the scenario number that is the cell link in this case.

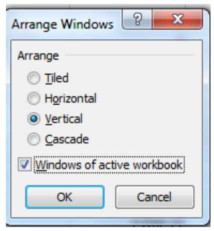


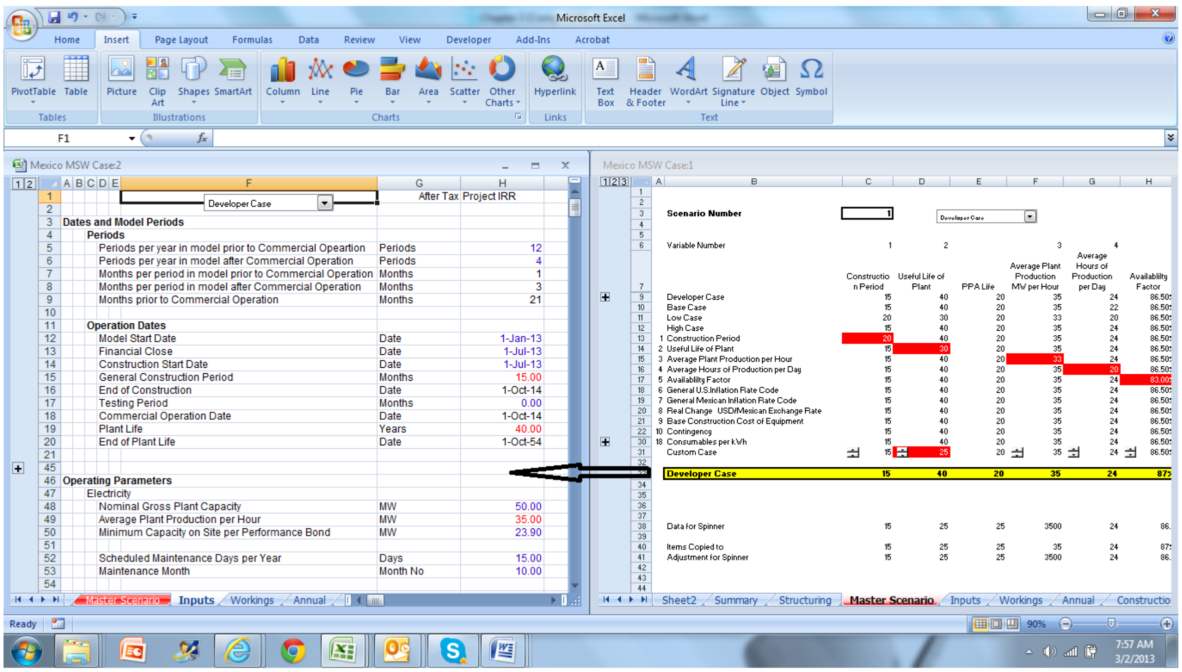


Step 5: Link Financial Model Inputs to Scenario Page

The only trick and somewhat awkward aspect of the scenario method using INDEX and DATA TABLE is that financial model inputs must be linked to assumption inputs in the master scenario page. Say that one of the inputs on the master scenario page is for the price of the product and many alternative prices are input for the different downside, upside, recession and other cases. Through using the INDEX function described above, one of the price inputs depending on the scenario selected is listed at the bottom of the paged depending on the scenario number. The price input from the last line (or the line to the right of the inputs if scenarios are set-up with columns instead of rows) must now drive the model instead of the input that was originally entered. After linking the model input to the master scenario page, you can change the scenario number either manually or using the dropdown box and the model input will change.

To make the tedious process of linking model inputs to the master scenario page less tedious, you can display both the model sheet and the master scenario sheet on the same screen. Then, you can go to the model inputs, enter an equal sign and then link the inputs to the INDEX line of the scenario sheet. To present two different sheets of the same workbook on the same page, go to the VIEW menu and first press new window. This opens two version of the same sheet. Next, go to the ARRANGE ALL option and select the windows of active workbook check box and the vertical option as shown in the diagram below. After you have two sheets open use the CNTL PAGEUP or CNTL PAGEDN so that one sheet shows the financial model and the second shows the master scenario page. Then you can work down the financial model inputs and link the various inputs to the INDEX line of the master scenario page. The process is illustrated on the figure below where the arrow indicates that the links are made from the master scenario page to the financial model – the numbers in the financial model must come from the scenario page and not the other way around. In the diagram below, the numbers that come from the financial model are shown in red which is the colour of the master scenario page.





For inputs such as those shown above where there is a single input in the financial model the process is straightforward. For other inputs that are structured as time series variables and have changing values over time, the same method can be applied. Recall that when time series variables are entered, you can use a scenario number to define different series of values over time. Where the inflation rate changes from one year to the next and you have a few inflation rate possibilities that change from year to year, the method described in Part 1 was to enter a scenario number and use the INDEX function much like the method described above. To select different time series variables, you can enter different code numbers for the inflation scenarios in the master scenario page (illustrated in the table above for U.S. general inflation and Mexican general inflation). Then, in the input section of the model where you enter the time series variables with a scenario code, link the scenario code to the master scenario page in the same way as for the other variables as illustrated below.

General Inflation and Exchange Rate Assumptions

Exchange Rates used in Construction Assumptions							
Base Date for Currency		1-Jan-13	1/7/2013				
Pesos/Dollar		13.10	12.74				
Pesos/ Euro		17.04	16.8				
Dólar/Euro		1.30	1.32				
General Inflation Rates							
Inflation Rate Scenario	Number	1.00					
		2013	2014	2015	2016	2017	2018
Mexico Inflation - Base Case	Annual Pct	3.80%	3.91%	3.65%	3.60%	3.75%	3.75%
Mexico Inflation - Low Case	Annual Pct	1.00%	1.00%	1.00%	1.00%	1.00%	1.00%
Mexico Inflation - High Case	Annual Pct	4.00%	4.00%	4.00%	4.00%	4.00%	4.00%
Mexico Inflation - Base Case	Annual Pct	3.80%	3.91%	3.65%	3.60%	3.75%	3.75%
USD Inflation - Base Case	Annual Pct	2.00%	2.00%	2.00%	2.00%	2.00%	2.00%
USD Inflation - Low Case	Annual Pct	0.50%	0.50%	0.50%	0.50%	0.50%	0.50%
USD Inflation - High Case	Annual Pct	2.50%	2.50%	2.50%	2.50%	2.50%	2.50%
USD Inflation - Base Case	Annual Pct	2.00%	2.00%	2.00%	2.00%	2.00%	2.00%
Peso/Dollar Real Devaluation	Factor	100%	100%	100%	100%	100%	100%

Comes from the master scenario page

Uses the INDEX function with the scenario number

If you use the method above, one minor problem is that key financial model inputs come from the master scenario page. If you would rather keep the inputs in the model and make the scenario analysis independent of the master scenario page you can use the VBA approach described below.

Step 6: Link a set of outputs alongside the scenario inputs

Once the INDEX function has been used to drive financial model inputs from the master scenario page, you can create a table that lists outputs from the model next to the various scenarios. The outputs should be placed one row above and one row to the right of the scenario numbers in order to allow construction of a data table. You should then enter a counter for the various scenarios one column to the left and one row below the output variables as illustrated below (the inputs for the various scenarios are hidden in the table). Once the output variables are structured in this manner, the data table tool can be used as described in the next step.

Scenario Number 1   Debt Leverage 70%

		Pre-Tax IRR USD	Pre-Tax Peso	Pre-Tax IRR Real	Equity IRR - USD	Equity IRR - Peso	Equity IRR- Real
		20.82%	25.26%	18.45%	22.18%	24.27%	19.79%
Developer Case	1	20.82%	25.26%	18.45%	22.18%	24.27%	19.79%
Base Case	2	20.45%	24.97%	18.09%	20.94%	23.01%	18.57%
Low Case	3	14.06%	16.10%	13.49%	10.62%	11.17%	10.07%
High Case	4	25.45%	30.42%	22.39%	46.23%	48.37%	42.67%
Construction Period	5	21.07%	25.55%	18.69%	24.02%	26.13%	21.59%
Useful Life of Plant	6	21.51%	26.08%	19.13%	24.32%	26.44%	21.89%
Average Plant Production per Hour	7	20.07%	24.42%	17.72%	20.38%	22.43%	18.02%
Average Hours of Production per Day	8	17.77%	21.83%	15.46%	15.92%	17.90%	13.65%
Availablilty Factor	9	20.51%	24.92%	18.15%	21.40%	23.47%	19.02%
General U.S.Inflation Rate Code	10	25.07%	28.29%	24.45%	33.73%	34.39%	33.06%
General Mexican Inflation Rate Code	11	21.57%	26.12%	19.19%	24.38%	26.50%	21.94%
Real Change USD/Mexican Exchange Rate	12	21.57%	26.12%	19.19%	24.38%	26.50%	21.94%
Base Construction Cost of Equipment	13	20.14%	24.31%	17.79%	20.09%	22.14%	17.74%
Contingency	14	20.80%	25.15%	18.44%	22.10%	24.19%	19.71%

These are variables that come from the financial model

Step 6: Create a one way table to list the results of multiple scenarios

The final step of creating a scenario analysis is to use the data table tool to list the results of each scenario. This involves the same techniques as described above for the break-even analysis where a one-way data table was introduced. Once the scenarios are set-up in the manner described above with a counter to the left and the outputs above, the data table is created by shading the area and then using the data, what-if and data table tool. The outputs are structured by first entering a counter for the scenario number and then linking output variables one row above and one row to the right of the scenario numbers in order to allow construction of the table. One the data table is shaded, the column input of the data table is assigned to the scenario number which cycles through each scenario and lists the output variables for the scenarios.

With the data table created alternative structuring variables can be used for the various sensitivity analyses along with the scenarios. The table above shows a spinner box next to the debt leverage. When the leverage is changed, all of the scenarios change along with the change in leverage. In this way one can test things such as how much debt a project can support before the DSCR, LLCR or the PLCR fall to a level of 1.0 in the downside case or the stress case.

Using VBA to Create a Scenario Analysis

Two problems with a the scenario process discussed above involve the fact that data tables are used which slow down the spreadsheet calculations (even if the calculation option without data tables is selected) and the fact that the financial model variables are linked to the master scenario page meaning that the input section of the financial model no longer contains inputs that can be adjusted manually without going to the master scenario page. Both of these problems can be solved through programming a macro. The first part of the macro assigns data to inputs in the financial model without linking the variables. The second part of the macro replaces the data table with a FOR NEXT loop as explained above for the break-even analysis.

Creating a macro that assigns variables from the INDEX line of the scenario page to the financial model inputs

To create VBA code that links the financial model inputs to the master scenario page one can create a macro with the following code (the range names such as C24 are replaced with general descriptions:

Range(“financial modelling inputs”) = Range(“Index formula for item in index function on scenario page”)

This code will put an input in the various financial model inputs, but will leave the cells with input data rather than be directly linked with an equal sign to the master scenario page. As with the other VBA discussed above it is important to name the ranges when creating a macro. In this case each input in the financial model that will be assigned a variable from the master scenario page should be assigned a range name (see the SHIFT, CNTL, F3 short-cut key above to do this.) For the INDEX range that will be used to transfer the variables, a range name can be assigned to the whole row. Once the range name is assigned, the CELLS function can be used to find the column of the range name to assign.

To illustrate the approach using VBA code and assume the price variable is being adjusted from the master scenario page. In the financial model, the range name is assumed to be named price and the row computed with the INDEX function in the master scenario page is named scenario. In this range,

assume that the price variable is the third variable. In this case, the VBA code would have the following form:

```
Range("Price") = Range("Scenario").Cells(3,1)
```

This code would have to be added for each variable that is changed by the master scenario page. The macro can be assigned to the dropdown box described above so that whenever items in the dropdown box are changed, the macro will assign the variables to the financial model.

The second VBA code that can be effective in developing scenario analysis is to use a macro that creates the data table. The method of using FOR NEXT loops and the CELLS function is described above in the break-even section.

Getting the Best of Both Worlds by Creating a Special Custom Scenario to Combine Scenario Analysis with Sensitivity Analysis

To create a combined scenario and sensitivity analysis, you can add an extra scenario below all of the rest of the scenarios. For variables in this special scenario, you can add spinner boxes or combo boxes. In the example above, the last column named the custom case is an example of this concept. For this scenario, each input that is not driven by an index number where the numbers change over time is assigned a spinner box. Those inputs that have changing values over time are assigned a code number are assigned a combo box. In this way one can choose the custom scenario as the scenario number and then change individual variables with spinner boxes or dropdown boxes.

There are two problems with creating this type of custom case. The first problem is that you would probably like to re-set the custom case to the base case or another case after playing around with the spinner boxes and combo boxes as you will soon forget value for the base case. The second problem is that when applying the spinner boxes, many of the inputs are expressed as percentages or as very large numbers which means that adjustments will have to be made. To solve this problem, you can create a little VBA code that re-sets the scenario to the base case or another case whenever the base case or the alternative is selected.

The mechanics of this process involve setting up a special row that is adjusted for problems with the spinner boxes in a separate row below the custom case. This row contains the data referred to in the spinner box such as 10 instead of 10% or 100 instead of 100,000. To convert this row back to percentages or large numbers, another row with multipliers can be used. In the above example with 10% and 100,000 the first multiplier would be 100 and the second multiplier would be 01. The multipliers have two functions. The first function is to multiply the data that is not adjusted so it can be used in the spinner box. The second function of the multipliers is to re-adjust the spinner boxes so that they can be used as inputs to transfer to the financial model. Once the multipliers are entered, the following step by step process can be used to re-set the custom scenario case.

Step 1: Begin recording a macro and copy the base case, the low case or any other case to a row a few lines below the INDEX function line that transfers the scenario variables to the financial model. This line (which may be titled raw data) cannot be used with spinner boxes if any of the variables are expressed as percentages or may have values larger than 30,000 (the upper limit of the spinner box) or have values that can change in other than whole numbers. After copying the scenario row (or column) to the raw data line, stop recording the macro.

Step 2: Enter the various multipliers in a row above or below the raw data row copied from above. If the variables are code numbers time series variables or are numbers between 10 and 30,000 that can be expressed as whole numbers without fractions then the multiplier can be 1.0. For other variables that are expressed as percentages, where a fraction is desired or where the number is larger than 30,000, a multiplier should be entered that is a multiple or a fraction of 10. An illustration of this process is shown below were the multipliers are 1.0 for variables such as the construction period, but are adjusted for percentages such as the contingency percent.

Scenario Number	30									
Custom Case										
	Construction Period	Operating Life	Code for Actual Capacity	Budget Contingency	Actual Contingency	Capacity Factor Code	Actual Unplanned Outage	Actual Fuel Price	Heat Rate	O&M Charge Escalation
Custom Case	38	25	Base Cas 1	0.0784	7.84%	Base Case	500	2.56	9000	Base Fixed C
Custom Case	38	25	1	0.0784	0.0784	1	500	2.5627	9000	1
Identifier for Code Numbers			Base Case			Base Case				Base Fixed O&M
Copy of Adjusted - Tied to Spinner Always Li	38	25	1	78.4	78.4	1	500	2562.7	9000	1
Including Formulas for Spinner - Adjusted	38	25	1	78.4	78.4	1	500	2562.7	9000	1
Multipplier	1	1	1	0.001	0.001	1	1	0.001	1	1
Raw Copy from Case	38	25	1	0.0784	0.0784	1	500	2.5627	9000	1

Step 3: Multiply or divide the row copied from the base case or another case by the multipliers in a third row. This row will have values that can be adjusted with either spinner boxed or combo boxes. In the above example, the multiplier for the contingency percentage is .001 and the raw data copied from the case is divided by the multiplier giving a whole number that can be adjusted with the spinner box. A similar process is used for fuel prices that should be entered with a fraction.

Step 4: Start recording a macro and copy the row that is computed from the multiplier by the raw data from the case and paste it in yet another row with fixed values rather than formulas. This is the row that will be used by all of the spinner boxes and the combo boxes for the custom case. It is necessary to copy and paste this row because the spinner box will fix values and you will surely want to copy another row to the raw data row to evaluate other cases later on. If you do not use this intermediate step, you can only perform the process after copying the base case one single time. If you would like to change the base case or run the custom case from another case such as the low case, then you must retain the flexibility where the raw data is adjusted by the formula without fixing the data and then fix the data in a second copy and paste macro. Once you have copied and pasted the row, stop recording the macro. For cosmetic purposes, after you have recorded the macro add a line in the VBA code to turn off the marking for the copied row using the formula:

```
Application.CutCopyMode = FALSE
```

Step 5: To create the spinner box, refer to the row in the last step that was copied and pasted to a fixed value. This value should be a whole number that can be adjusted by the spinner box and later adjusted through dividing or multiplying the number by the multiplier factor. The line could be called the spinner box input row and it should have fixed values rather than calculations.

Step 6: Once you have computed the adjusted fixed row – the raw data adjusted for the multiplier – you can establish the custom scenario line that will be used in the model. The custom case scenario uses the spinner box line that you created containing fixed adjusted values and then

re-adjusts this line using the multiplier factors. For example, if the input is an interest rate of 5.2%, then the value expressed as a decimal is copied to the raw data row --.052. Next the row data row could be multiplied by 1000 yielding 52. This value of 52 is copied to a fixed spinner row which is attached to the spinner box. Finally, the number in the custom row of the scenario table and used to transfer data to the model is divided by 1000.

Step 7: Put all of the things from the above steps together in a macro an use range names for the cases that are copied and the lines that are copied (for example, the name the base case row, the multiplied row and row used for the spinner box). The macro will be attached to the dropdown box so that the basis for the custom case can be the base case, the low case or any other case. To effectively program a macro that copies different cases depending on the scenario chosen, you can use the SELECT CASE command followed by the END SELECT statement in the VBA code. This method will allow you to copy different cases to the raw data described in the first step depending on which scenario is selected. Using the SELECT CASE statement, you can apply a different copy and paste routine depending on the scenario number selected. To use the SELECT CASE you include the case that will drive the different code – in this case it is the scenario number. In the VBA code you would write:

SELECT CASE RANGE(“Scenario number”)

Below this command you would enter CASE 1: and CASE 2: and so forth depending on how many cases can be copied to the raw data line and the used in the custom case which is driven by the number of scenarios. Underneath CASE 1: you then put the copy and paste statement that copies the base case to the raw data line below the INDEX function as illustrated below:

CASE 1:
Copy base case to raw data row
CASE 2:
Copy low case to raw data row
…..

Once you are finished with the copy and paste functions in the SELECT CASE analysis, you must write the END SELECT statement. Then, below the end select statement make another copy and paste routine. This part of the macro copies the line computed from the raw data adjusted for the multiplier to yet another line that is fixed. This fixed line is used by the spinner box. Finally, if you do not want the copy and paste functions to apply to selections from the drop down box other than the base case, the low case and the high case, you can use the following code:

Case Is > 3
Exit Sub

Chapter 19: Dissecting and Presenting the Effects of Different Variables in Scenario Analysis through Generating Tornado and Spider Diagrams to Impress People

In scenario analysis, there is often a very large difference between the base case and the downside case. For example, the base case may have a 12% IRR while the downside case could have a 2% IRR. To make sense of the scenario analysis it is useful to understand which of the input variables are most instrumental and least instrumental in causing the IRR's to change. This analysis is both useful in risk analysis and it can help to evaluate potential errors in model assumptions. When isolating the effect of individual assumptions on the difference between scenario outcomes you make a number of different cases where the base case is adjusted by one variable at time. Once you have created these sensitivity cases you can present this analysis in two different forms. A tornado diagram shows which variable has the largest effect and a spider diagram presents the increments of changes in different variables from the higher case to the low case.

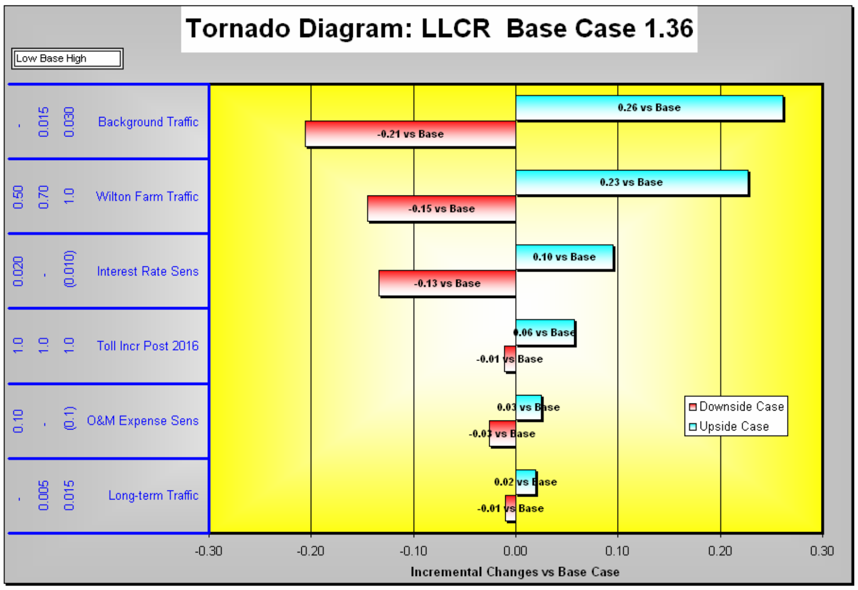
This chapter describes why you may want to present a tornado diagram or a spider diagram in risk analysis and it explains how to build the graphs from the scenario analysis discussed above. The next paragraph addresses two different ways to create a tornado diagram and the subsequent section discusses spider diagrams.

Tornado Diagrams that Display which Variables have the Largest Effect on Value and Which Variables have the Least Effect on an Output Variable in the Shape of a Tornado

The scenario analysis discussed above did not show which of the input variables are most important and which are not very important with respect to one of the output variables. This can be accomplished with a tornado diagram. A tornado diagram gets its name because it is supposed to look like the whirlwind created by a tornado which is larger in size on the top of a graph and smaller on the bottom. In a tornado diagram, input variables which have the highest effect on a selected output variable are shown at the top of the graph while variables which have a smaller impact are shown at the bottom. Through presenting sensitivity analysis in this manner, a tornado diagram can be used as a tool to determine which variables are most important in a sensitivity analysis and should be the focus of more study and due diligence. In addition, this type of presentation can be used to illustrate whether an investment has relatively more upside risk or downside exposure – if the tornado is bending toward the upside case or the downside case as tornadoes sometimes do. Thirdly, a tornado diagram can be used as a tool to test the mechanics of a financial model. If the variables which have the highest effect are not intuitive, one should go back to the financial model and carefully examine the formulas related to the non-intuitive variable. Finally, a tornado diagram is simply a way to show how smart you are with excel and impress or confuse people who are reading the model.

The figure below illustrates how a tornado diagram can be interpreted in risk analysis. Data is entered for a number of variables in base case, downside case and upside case scenarios just like the scenario analysis above. While the format is similar to the scenario analysis, the downside and upside cases do not necessarily have to represent a consistent outlook, meaning for example that if sales decrease, than operating costs may also decrease in a scenario analysis. When reviewing the effect of each variable, the downside case simply represents a low range outlook with respect to each variable. In the figure below, the data entered – on a judgmental basis -- for base case, downside case and upside cases are shown on the left hand side of the graph. Given these inputs, each bar on the diagram shows how the variation in the downside case from the base case, or the upside value relative to the base case affects the variable in question (shown on the top of the graph.) In the diagram below, the variable being studied is the loan life coverage ratio ("LLCR") for a toll road project financing. The tornado diagram shows that given the input assumptions, the variable which has the highest effect on the LLCR is the background traffic. This variable causes the LLCR to fall by .21x, when all other assumptions use base case values, when the traffic growth is 0% rather than 1.5% (the downside case and base case

assumptions shown on the left of the graph.) Similarly, if the background traffic increases by 3% per year, the LLCR increases by .26x relative to the base case. On the other hand, variables such as the operation and maintenance expense and the long-term traffic projections have a relatively small effect on the LLCR, meaning that if the LLCR is an important measure for analysis, one would not have to worry a great deal about these variables as long as the judgmental downside and upside assumptions are reasonable.



The Easy Way to Create a Tornado Diagram by Extending Scenario Analysis

A relatively simple way to make a tornado diagram similar to the one displayed above is to begin with the scenario analysis discussed in the last chapter, add two new scenarios for each input variable – named sensitivity scenarios -- and then create a data table summarising all of the scenarios. Once a large data table is constructed with the additional scenarios, information from the data table output is used to present a graph like the one above. As with the scenario analysis, the entire process can be completed with a data table and the INDEX function. Various detailed steps to construct the scenario analysis are listed below:

Step 1: After entering the basic scenarios (the base case, the low case etc.), you can enter a new scenario for each variable through transposing the row names (or the column names if the variables are entered in different rows.) These new scenarios are labelled sensitivity cases in the discussion below. For example, if one of the variables represents prices, another represents quantities and a third represents interest rates, then these variable names are presented as sensitivity scenarios below the low case. The scenario list illustrated in the last chapter is an example of such a scenario list. To transform the variable names you can use the TRANSPOSE function introduced in Part 1 (recall that you use SHIFT, CNTL, ENTER rather than the enter key).

Step 2: After listing the different sensitivity scenario cases, link all of the values of the base case to every single new sensitivity case lines that you just listed for the variables. To do this you can link one of the variables (with the equal sign) and then press the F4 key twice to change the relative references and fix the dollar sign only on the row or column number. When you have created the scenario for one case, copy it for the entire block of remaining data.

Step 3; Change the value of the variable that is listed in the sensitivity case to the low case (for example, change the price variable from the base case to the low case). This results in a diagonal pattern where all variables have the same value as the base case except one single value for each sensitivity case scenario. Creating these multiple scenarios allows you to isolate the effect the downside value for each value. To illustrate which values in a scenario are changed from the base case you can use the conditional formatting option with the formula option. To implement this approach where the format depends on the equality of two values, enter the current cell in the conditional formatting slot without any fixed references (without the dollar signs). Then test whether the variable is not equal to the base case by using the less than and greater than sign, <>. When testing the value relative to the base case, use the fixed reference on the row number for the base case so the conditional formatting will work across all of the scenarios.

Step 4: After adjusting the individual variables in to the downside case, follow the same process for the upside case meaning that you transpose the titles again, link the base case values to all of the assumptions and then only change the input value to the upside case for the scenario in question. For example, after you include an additional line for the upside case price scenario, you then link the price variable to the upside case value. An illustration of setting-up the downside and upside scenarios is shown below. Note that the phrase "low case" and "high case" is attached to the various scenario titles through appending the TRANSPOSE function.

Scenario number	1	Base case
Synergy Sensitivity	100%	
Premium	20%	

DCF Analysis

	Sales Growth	COGS margin	Sales/AR	Inventory/ COGS	Capex/sales	AP/Expenses		Equity IRR	Growth rate method	Value driver method
								21.44%	28.99	25.98
Base case	1	1	8.40%	5.60%	1	5.60%	1.00	19.29%	34.54	30.16
Low case	2	2	10.00%	10.00%	2	3.00%	2.00	11.33%	18.85	15.10
High case	3	3	7.00%	5.00%	3	7.00%	3.00	23.92%	50.87	41.53
Sales Growth Low Case	2	1	8.40%	5.60%	1	5.60%	4.00	13.69%	22.35	18.36
COGS margin Low Case	1	2	8.40%	5.60%	1	5.60%	5.00	17.71%	29.99	26.52
Sales/AR Low Case	1	1	10.00%	5.60%	1	5.60%	6.00	19.06%	35.36	30.57
Inventory/ COGS Low Case	1	1	8.40%	10.00%	1	5.60%	7.00	19.29%	34.54	30.16
Capex/sales Low Case	1	1	8.40%	5.60%	2	5.60%	8.00	19.12%	34.43	29.69
AP/Expenses Low Case	1	1	8.40%	5.60%	1	3.00%	9.00	19.29%	34.54	30.16
Sales Growth High Case	3	1	8.40%	5.60%	1	5.60%	10.00	22.84%	44.41	40.02
COGS margin High Case	1	3	8.40%	5.60%	1	5.60%	11.00	18.85%	32.05	28.31
Sales/AR High Case	1	1	7.00%	5.60%	1	5.60%	12.00	19.48%	33.82	29.79
Inventory/ COGS High Case	1	1	8.40%	5.00%	1	5.60%	13.00	19.29%	34.54	30.16
Capex/sales High Case	1	1	8.40%	5.60%	3	5.60%	14.00	20.55%	43.55	33.85
AP/Expenses High Case	1	1	8.40%	5.60%	1	7.00%	15.00	19.29%	34.54	30.16
Custom Case	Base sales	Base margin	8.40%	5.60%	Base Capital E:	5.60%	16.00	19.29%	34.54	30.16
Base case	1	1	8.40%	5.60%	1	5.60%				

Step 5: Once data for the low case and the high case sensitivity cases are entered adjust the INDEX function and the DATA table to reflect the additional cases. To adjust the INDEX function, simply include the additional rows in INDEX calculation. Also you must add the new sensitivity case numbers in the column to the right of the data table. Recall that the creating the data table is developed by shading the lengthened data table and attaching the column input to the scenario number. As the numbers from the INDEX row are already attached to the financial model inputs, the data table should work fine.

Step 6: To create the tornado diagram from this process, pick one of the output values from the data table (that should now show a different value in each scenario, including the sensitivity scenarios). To evaluate the effect of each variable separately from the others, the output values produced by the data table for the various sensitivity cases relative to the base case value in the data table can be compared. To make this evaluation, subtract values in the data table for each sensitivity case from the base case value in the table. When computing the change of the sensitivity case relative to the base case, create a list of the variables that are changed and next to this list place the change in the upside case relative to the base case and then the downside case relative to the base case as illustrated in the table below. Calculations for these differences can be made by fixing the base case cell in the data table with the F4 key and then subtracting each of the different scenarios one at a time. This table could be placed below the master data scenario.

	Low v Base	High v Base
Sales Growth	-6.23%	3.88%
COGS margin	-1.72%	-0.49%
Sales/AR	-0.24%	0.21%
Inventory/ COGS	-0.34%	0.05%
Capex/sales	-0.16%	1.37%
AP/Expenses	-0.33%	0.18%

Step 7: With a table shown above, you can make a graph that demonstrates which variables have the highest effect on the output variable and which have the least effect on the variable. Simply shade the table above and use the F11 key to create a graph and then change the chart type to a bar chart. The only issue with this graph is that it is not yet sorted for presentation. The graph is illustrated below.

Step 8: To sort the data from highest to lowest on the graph, follow the steps beginning with step 10 in the discussion below.

How you can Create a Tornado Diagram Using a Two Way Data Table

Mechanical implementation of a tornado diagram can also be accomplished with a two way data table where each variable is changed one at a time with a row input for the variable as well as the scenario number for the base case, low case or high case as was described above for a one way table with a scenario analysis. Using this method you do not have to re-enter each scenario over and over again but it is difficult to present the scenario analysis along with the tornado analysis without adding an option button. Further, this is very similar to the method that is used to create spider diagrams (another impressive gimmick) described in the next section. As with the above method the process involves isolating the effect of one variable while all of the other variables are held at the base case value. In the tornado diagram graph shown, the fist bar for the background traffic is changed from the base case to

the low case to the high case while all of the other variables are held at the base case value. When using the two way data table to compute a tornado diagram, the base case for any variable is the starting point and the output is always re-set to the base case value as all of the other variables also remain at the base case number. To create a tornado using this process you can follow the following steps:

1. Set up a scenario analysis using the INDEX function together with the data table as described above for the normal scenario analysis, where the column is used in the INDEX together with the scenario number. This accomplishes two things. First, it allows you to maintain the scenario analysis. Second it provides a basis for creating data for the tornado diagram.

2. Add a form control that allows one to select between a scenario analysis and a tornado diagram as the link will either come from the INDEX from the scenario analysis or a second adjusted line for the tornado diagram. (You can use option buttons where you insert a button and then select a cell link. The option buttons work well with the CHOOSE function in excel.)

3. To create a two-way data table rather than a one way table, add an input for a variable number as well as the scenario number. The variable number could be entered near the scenario number that was used in the scenario analysis. In addition to this variable number that is analogous to the scenario number, enter a number associated with each variable in an analogous manner to the way you entered the scenario number to the left of the data table in developing the scenario analysis. Entering the variable number and the codes for each variable is illustrated in the diagram below.

4. Below the INDEX computation that has been discussed at length for the basic scenario analysis, add a new row that includes a true/false test for whether the variable in the table is the same as the variable number input. If for example, there are six variables as illustrated in the diagram above, there should be a list of numbers from one to six above the variables. The test evaluates whether the input number for the variable number equals the number associated with the variable. There is only one value that is TRUE for the list of variables. The formula for the test should use the F4 key for the variable number and compare this to each formula as illustrated below:

    =Variable Number Input (Fixed with F4) = Number Associated with Each Variable

5. Below the test variable create another row. This row uses an IF function with the test TRUE/FALSE variable. If the test variable is TRUE, then use the value from the INDEX function discussed in step 3 above. If the variable is false, then use the base case value. Using this approach, every variable remains at the base case value except the variable which equals the variable number input. This is the trick to make the two-way table be able to present the how each variable changes one at a time in isolation. The IF test has the form illustrated below:

    IF(TRUE/FALSE Test, INDEX from normal Scenario in Step 3, Base Case from Scenario Table)

6. In order to make the model operate with either the scenario analysis or a tornado diagram, you have to enter yet another row. This row simply selects whether the normal index function is used to drive the inputs in the model or whether the IF statement above that sets all variables to the base case except for the single variable in question is used. To make this selection you can use the CHOOSE function together with the cell link described in step 2 (and shown in the diagram below) to select either the INDEX value – for the scenario option, or the IF function –

for the tornado option. As described in detail above for the scenario analysis, the values in this final row should now drive the financial model and the model assumptions should be linked to this line. An illustration of this part of setting up a tornado diagram is illustrated in the excerpt below. The last line in the table is used in the input section of the financial model.

Scenario Number	2			
Variable Number	2	Scenario	2	Run Table
Shares Issued	5,000.00			
Premium	40%	Tornado		

	Trafic per Store	Margin per Store	Cap Exp per New Store	Growth Rate in Stores	Cost Synergy Percent		IRR on Equity - PE Ratio	IRR on Equity - M/B Ratio	Overall IRR on Free Cash Flow	5-Year Average Accretion	DCF Value - Equity
							12.78%	13.49%	9.36%	10.12%	17,769.60
	1	2	3	4	5						
Base Case	230,000.0	30.0%	850,000	4.50%	10.00%	1	13.3%	14.0%	9.8%	12.4%	19,049.43
Low Case	210,000.0	29.0%	930,000	3.00%	0.00%	2	5.8%	7.3%	5.1%	-13.0%	7,479.95
High Case	250,000.0	33.0%	800,000	6.00%	15.00%	3	18.3%	18.7%	14.1%	29.4%	32,683.48
Worst Case	210,000.0	28.0%	930,000	2.00%	0.00%	4	3.9%	6.0%	4.4%	-14.3%	6,728.93
Index Function	210,000.0	29.0%	930,000	3.00%	0.00%	=INDEX(H12:H15,Scenario_Number)					
Test	FALSE	TRUE	FALSE	FALSE	FALSE	=$D$4=H11					
If Test	230,000.0	29.0%	850,000	4.50%	10.00%	=IF(H18,H17,H12)					
Low Case	230,000.0	29.0%	850,000	4.50%	10.00%	=CHOOSE($I$4,H17,H19)					

7. After developing scenario inputs, the variable number, the test TRUE/FALSE variable, the if test and the CHOOSE function as shown above and after linking the input variables in the model to the CHOOSE function instead of the INDEX function row, you are ready to set up a two way data table where one of the variables represents the variable number and a second variable represents the scenario number for the base case, low case and high case. The form of this two way data table is illustrated below:

		Variable Number				
		1	2	3	4	5
Sc Num	1					
	2					
	3					

8. In the above table, the first row represents the base case, the second is the low case and the third is the high case. In terms of columns, each column contains the value of the selected output variable if the input variable is changed and every other variable remains at the base case value. Once the table is set-up, run either a two way table or if you want to be a little fancier, write some VBA code that does the same thing.

9. For the first row, the base case output is simply repeated for in each case. Once the table is computed, the base case versus the low case and the base case versus the high case can be computed by simply subtracting the low case row from the base case row and by subtracting the high case row from the base case row. This presents the impacts of different variables which is the basis for the tornado diagram. All that is left to do is to present the data. An illustration of this step is shown below. You can see that the variable with the largest effect if variable number 5 and the second largest is variable number 2 with variable number 3 coming in last place.

		Variable Number				
		1	2	3	4	5
Sc Num	1	13.3%	13.3%	13.3%	13.3%	13.3%
	2	11.4%	12.8%	12.9%	12.8%	10.8%
	3	14.9%	14.8%	13.6%	13.9%	14.5%
Low vs Base		-1.9%	-0.6%	-0.5%	-0.5%	-2.6%
High vs Base		1.6%	1.5%	0.3%	0.5%	1.2%
Absolute Value		3.5%	2.1%	0.8%	1.0%	3.7%

10. In presenting the data, the idea of a tornado diagram is to sort the absolute value of the variables from the smallest to the largest to the smallest so the graph will have the shape of a tornado as pictured above. Unfortunately, use of the sort function in excel is not helpful in this instance for a variety of reasons. First you probably do not want to change the order of all the variables in the scenario table. Second and more importantly, if you have used the data table function, you cannot sort items in the data table.

11. Instead of using the excel sort function, you can use the SMALL function along with the MATCH and the INDEX function where the original values will be left alone and you do the sort process somewhere else. After using the SMALL function to sort the variables, an essential step is to create a single variable that will be a sort key used for tabulating the variables in a sorted order using the MATCH function. Once the sort key is established, you can use the INDEX function to present selected variables in sorted order and make your nice graph.

12. To complete this process, make a new column or row next to or below the sum of the absolute value of the base versus low and base versus high differences. Then, enter a number for each of the variables (as the SMALL function gives you the first smallest value, the second smallest value and so forth). Then enter the SMALL as illustrated below:

    SMALL(Row or column to Sort (Absolute Value from Table Above), Variable Number)

13. Once you have sorted the absolute values using the method above, you have the sorted values but you do not have the names associated with each of the sorted values. To retrieve the names associated with each of the values and to retrieve the base versus low and the base versus high, you can make a counter or index variable that will move the titles and the other values into the correct sorted slots. To do this the most important step is to make a MATCH function that produces the original row or column number of the sorted variables, which could be called the sort key. This sort key that has the original row or column number for the sorted data is essential because it is the basis for the INDEX function to produce the original names and data to graph. When computing the sort key, use the MATCH function so as to create an exact match. To do this enter a zero as the match type as illustrated below:

    MATCH(Sorted Single Value, Unsorted Values -- Fixed, 0)

14. Sometimes two or more of the sorted variable have the same value. In this case you can add the RAND()*.000001 to the variable to be sorted so that there is a unique value and then apply the same process.

15. To create a graph, you will need the title of each variable as the x-axis and the low versus the base and the high versus the base. To set-up the data for making a graph, use the INDEX function along with the sort key. For example, to find the title, enter the original titles (fixed with

the F4 key) in the index command and use the sort key as the column number (you do not need a row number) as illustrated below:

INDEX(Original Series of Titles – Fixed, Sort Key)

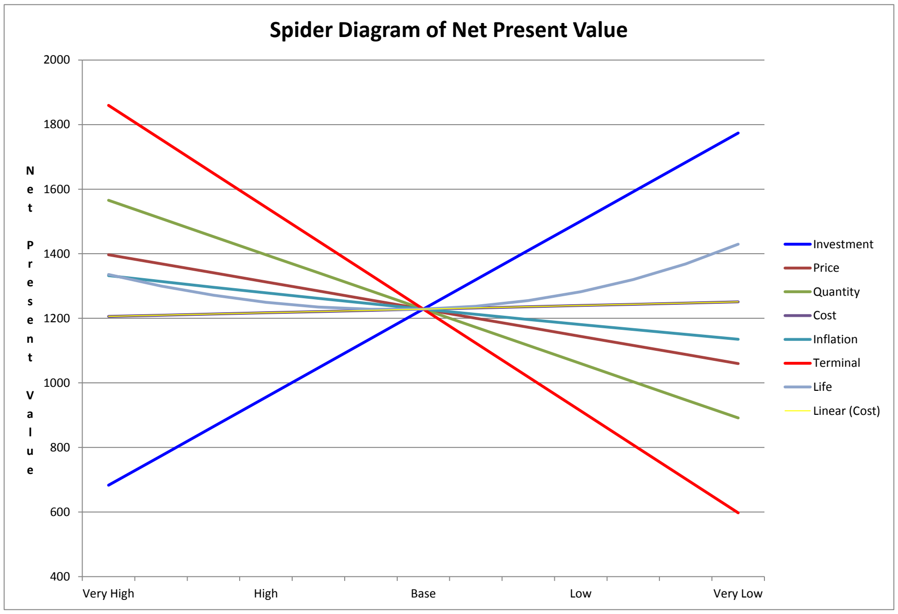
	Cap Exp per New Store	Growth Rate in Stores	Margin per Store	Traffic per Store	Cost Synergy Percent	=INDEX($D$10:$H$10,H37)
Low vs Base	-0.49%	-0.51%	-0.57%	-1.92%	-2.56%	=INDEX($D$32:$H$32,H37)
High vs Base	0.27%	0.51%	1.50%	1.57%	1.16%	=INDEX($D$33:$H$33,H37)

16. Graph the sorted variables using the F11 function and change the chart type to a stacked bar chart (make sure that there is no caption next to the titles so excel will know that this is the x-axis).

Spider Diagrams that Illustrate how Each Input Variable Affects and Output Variable and Presents Results that Look Something like a Spider Web

Some people like to see a where percentage variation is shown on the x-axis of a graph and one of the output variables is presented on the y-axis. The graph is presented for many of the variables and one can see how the output variable reacts to changes in the various input variables. If the variable is the IRR then the price variable should have an upward sloping line because price increases increase their while the cost variable would have a downward slope. Variables that have a steeper slope have more of an effect on the output variable while variables that do not have much of an effect should have a relatively flat line. While this kind of diagram may have some minor interest, the big problem is that each variable may not have the same potential variation. Some variables may have a range of only 5% while others may have a range of 200%.

The graph below illustrates the format of a spider diagram where various items that affect the net present value of an investment are shown with ranges in their possible values. Instead of showing simple percentages on the x-axis, the x-axis displays potential values from very low to very high. Here, a judgment is made as to how much a potential variable can change from the base case in terms of percentage changes. The amount by which a variable could change is similar to the idea of volatility discussed in the next few chapters and is a transition to the stochastic risk analysis methods. On the graph below, the amount spent on the investment has the largest negative effect on the value as very low values lead to a high net present value. On the other hand, the terminal value has the opposite effect as higher terminal values lead to a higher net present value. The real question in this analysis is to determine how large the possible percentage variation (or the volatility) is for each of the variables.



How to Create a Spider Diagrams using a Two-way Data Table

The mechanics of a spider diagram are similar to the construction of a two way data table that was used to create the tornado diagram above. Indeed, the graph above is created directly by graphing the two way data table. The two way data table is generated by revising the financial model so that sensitivities in terms of percentage variation can be accepted; adding a row number associated with the different variables; linking the row number to the row of sensitivity factors that are used in the model inputs; creating a two way data table with the variable row number and the scenario number; and, finally producing the two-way data table that can be graphed. The step by step process for creating the spider diagram using this process is explained below:

Step 1: Set-up a table much like the scenario table discussed in the last chapter with the variable names on each column (you could set-up the table with variable names on each row, but the discussion below will assume different variables across different columns.) When setting up the table, enter percentages in each row rather than the direct values of the variables. For example, the first row may contain 10% across the whole row and the second row could contain 20% across all of the columns. Include enough rows so that the percentages eventually increase to well above 100%. Somewhere near the middle of the table a row with 100% across all of the columns should be included. These percentage changes will drive the output variable changes in the spider diagram. You can think of each of the percentage changes as different scenarios in a scenario table above. The trick is to make each variable change by one of percentage figures on a separate basis.

Step 2: Revise the model so that sensitivity factors can be accepted. This means that a factor should be added to each line number in a model and the calculations in the model should be adjusted to accept different values of the factor. For example, a factor could be inserted just below the price variable where changing the factor changes the revenues.

Step 3: Create a scenario number and a variable number in the same way that you entered the variable and scenario number for the analysis of a tornado diagram. As with the two-way table for the scenario diagram, enter a line number below the last row of the table using the INDEX function and the scenario number. Then, add a line that includes a TRUE/FALSE test derived from the variable number and finally, add and IF statement that applies either the number in the INDEX row or 100%, depending on the test variable. An example of inputs for a spider diagram is shown on the table below.

Scenario No	6							
Variable No	1							
				Variable Number				
		1	2	3	4	5	6	7
		Moisture Content	Heat Content	Heat Rate Sensivity	Construction Cost/Ton/Day	Capacity Factor	Tipping Fee	O&M Cost
	1	70.00%	50.00%	50.00%	50.00%	50.00%	50.00%	50.00%
	2	75.00%	60.00%	60.00%	60.00%	60.00%	60.00%	60.00%
	3	80.00%	70.00%	70.00%	70.00%	70.00%	70.00%	70.00%
	4	85.00%	80.00%	80.00%	80.00%	80.00%	80.00%	80.00%
Scenario Num	5	95.00%	90.00%	90.00%	90.00%	90.00%	90.00%	90.00%
	6	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%
	7	105.00%	110.00%	110.00%	110.00%	110.00%	110.00%	110.00%
	8	110.00%	120.00%	120.00%	120.00%	120.00%	120.00%	120.00%
	9	115.00%	130.00%	130.00%	130.00%	130.00%	130.00%	130.00%
	10	120.00%	140.00%	140.00%	140.00%	140.00%	140.00%	140.00%
	11	125.00%	150.00%	150.00%	150.00%	150.00%	150.00%	150.00%
Index Function		100%	100%	100%	100%	100%	100%	100%
Test on Variable Number		TRUE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE
Sensitivity Applied in Model		100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%

Step 4: Link the row created with the IF test discussed above to the financial model where inputs are accepted for the various percentage factors as described in step 2 above. As in for all of the different situations, the direction of the link is from the scenario or spider table towards the financial model.

Step 5: Create a two-way data table through entering the variable numbers on different columns and the variable numbers across the rows is explained above for the scenario analysis. Then pick one of the output variables from the financial model and place it in the upper left hand cell of the table and create the data table. As with the tornado analysis discussed above, the row input cell is the variable number and the column input cell is the scenario number. Once you have made the data table, add titles to the rows and columns. In the example above, the titles for the row are the variable names and the titles for the columns are the percentage changes.

Step 6: Use the data table to construct a line graph. To adjust the data table to make a graph, you can group (temporarily hide) the scenario numbers and the variable numbers that are the basis for the row and column inputs of the data table. This is accomplished by shading the single row and the single column and then using the SHIFT, ALT, → short-cut key.

Step 7: Adjust the inputs to reflect different potential variations in the variables beginning with the 100% line and then moving up or down by different percentages. For example, the price may vary by 20% to the upside and 10% to the downside, while the CGS percent may vary by 5% to both the upside and the downside. If there are five rows to the upside and five rows to the downside, the

percent change in each price increment is 20% divided by 5 or 4% while percent change in the downside case 10% by 5 or 2%. In the case of the CGS percent the upside and the downside percent are increased and decreased by 5% divided by 5 or 1%. When labelling the columns, you can use labels such as the high case, the very high case and the high range case. Everything else is similar to the table where all variables increase or decrease by the same percentage as in the table shown above. The process of adjusting ranges in this manner illustrates that variables do not have the same volatility; that there are lower or upper boundaries on variables and that the upside may not be the converse of the downside case.

Chapter 20: Advantages and Disadvantages of Adding Probabilistic Risk Analysis and Time Series Equations to Financial Models Rather than Using Judgmental Assessment of Alternative Scenarios

The remainder of the risk analysis discussion covers application of mathematics and probability analysis to risk assessment. The principal difference in using statistical analysis to gauge risk as compared to the above methods is that the mathematics and probability analysis supposedly does not depend in on the subjective opinions of people as to the assumption of how high or how low some variable may become in the future. These opinions as to the potential upside or downside direction of key assumptions used in scenario analysis above can be biased upwards by managers having a favourable attitude towards an investment concept. The judgment can also be biased downwards if people generally disagree with a business idea or as happened in the dark days of October and November 2008 after the financial crisis, had such a pessimistic outlook on the future that they would not even make a loan to corporations like Siemens. Such human biases can theoretically be avoided by using cold hard statistics from historic data or from general market expectations to come up with the range in values for the upside and downside case variables. In fancier language, the process of using statistical analysis rather than judgment to measuring risk can be termed applying a stochastic diffusion processes to prediction of economic variables. Thus, one way to think about this stochastic analysis is that ranges in variables for sensitivity analysis, break-even analysis and scenario analysis come from objective statistical analysis of historic data rather than from subjective judgment.

In addition to providing an objective basis for ranges in variables, applying probabilistic analysis in financial models allows you to answer questions that can not be resolved with sensitivity analysis, break-even analysis or scenario analysis. Examples of such questions that can be answered with mathematical risk analysis include what is the probability of the IRR to equity and different debt tranches being below the risk free rate; what is the required credit spread on alternative types of debt issues; what is the probability of achieving different levels of earnings per share; and what is the chance of financial ratios falling to levels below those of investment grade companies.

Some people reading the above paragraphs may already be red in the face about these supposed benefits of mathematical risk analysis which ultimately pretends that you can make a business into a mathematical equation with a probability distribution. A few of the problems in attempting to apply probability analysis in financial models include:

Problem One: History used in establishing probability distributions is often a very poor predictor of future economic variables that are driven in part by human behaviour. Economic variables such as price and demand growth are distinguished from physical variables such as wind speed, reserves of oil in a field or hydro conditions that do not depend on the whims of human beings. Prices and volumes can experience sudden jumps or falls

in prices when market structure changes or when surplus capacity increases above a certain level.

Problem Two: Parameters required for implementing probability analysis in financial models -- especially mean reversion, correlation, price trends, boundary levels, and jumps – are often very difficult if not impossible to estimate from historic data. However in many applications the mean reversion parameter has a more important effect on the outcome of the probability analysis than the volatility estimate.

Problem Three: If prices or other variables follow cyclical patterns with mean reversion, then an estimate of the long-run price must be included in the analysis. When parameters are included for volatility in long-term equilibrium prices or changes in volatility itself the whole process seems to boil down to a whole lot of random numbers.

Problem Four: The implicit or explicit assumption that rates of changes in variables follow a normal distribution is generally not valid which renders the whole process biased if not useless, especially when focusing on the tail end of distributions in estimating credit losses. Furthermore, attempting to incorporate alternative distributions in a financial model requires added parameters which can be almost impossible to estimate.

Problem Five: Even with efficient software and fast computers, the process of implementing Monte Carlo simulation into large financial models can be time consuming and very tedious.

Given these problems with attempting to convert a business into a mathematical equation, you may wonder why so much of the risk analysis discussion is devoted to the subject of stochastic modelling. The general idea behind describing the mathematical approach to risk analysis in some detail is not necessarily to advocate the approach or to suggest that you should go out and immediately apply time series and Monte Carlo simulation. Instead, reasons for becoming familiar stochastic risk analysis techniques include: (1) making sure that you will not be intimidated or overly impressed when presentations are made using the approaches; (2) encouraging you to fully understand the flaws as well as the benefits of the stochastic modelling techniques; (3) explaining how the mathematical techniques can be combined with business judgment to make the process applicable in some selected practical situations; and (4) using the stochastic representations of economic variables as a framework to think about how differences in the risk structure of different companies or projects. Finally, various problems in finance and economics that do not directly use Monte Carlo simulation involve statistical analysis. In renewable energy analysis, electricity production is represented by probabilities associated with wind speeds, solar irradiation and hydro production. Similar estimates are made for oil and gas reserves and traffic studies. The probabilistic aspects of time series and Monte Carlo simulation are directly related to other uses of statistical methods in risk analysis.

The fourth point above about thinking how key variables can move is worthy of a bit more elaboration. To implement a time series model for a variable, one must come up statistical parameters that represent long-term trends, variations around the trend, eventual reversion to long-run levels, lower and upper boundary values, possible sudden moves, and the relationship between the variable and other variables. Computing these parameters and -- more importantly -- thinking carefully about them forces you to consider how key variables can potentially move in the future. For instance, in establishing downside case and upside case assumptions one should think about what is the lower limit, how much can the variable move in a year, will the variable move back to a long-term equilibrium level, what is that equilibrium level, can a “perfect storm” cause a dramatic change in the variable, and how does the variable move with other variables. If you do attempt to implement mathematical risk analysis into a

financial model, thinking about the underlying economics is also important from the stochastic modelling perspective. When coming up with parameters in time series equations, you need to understand the underlying economic factors which drive parameters such as volatility, mean reversion parameters so that you can better understand the real sources of risk in an investment.

What the Terms Time Series Equations, Volatility, Mean Reversion and Monte Carlo Simulation Really Mean

A few terms that are used in creating probabilistic equations from financial models include time series equations, volatility, normal distributions, mean reversion and Monte Carlo simulation. The manner in which probability analysis can be applied in financial models is through expressing variables in terms of time series equations. Time series equations are mathematical formulas that describe an economic variable in terms of its potential dispersion as well as the expected level of the variable. For example, the formula for price in a financial model described in part one may be $P_t = P_{t-1} \times 1.1$. This simply means that the current year's price is last year's price increasing at a rate of 10%. An equation such as this is a deterministic equation. If the equation was a time series model, then the next year price would be modelled with a possible dispersion – say the price may increase by 20%, by 10%, by 0% or by percentages within the range. The expected value is still a 10% increase, but the time series equation includes probabilities of different values within the range. The manner in which the price can vary around the expected value of an increase of 10% is defined by a volatility parameter. If the volatility is close to zero, there is little chance of the price moving by much more or much less than 10%. If the volatility parameter is high, the price can move in a wide range.

Using time series equations rather deterministic equations in a model allows one to project ranges in values rather than only focusing only on one case at a time. This also means that all of the outputs of the model such as the IRR, the DSCR, the amount of debt that is repaid and any other variable to be expressed as a probability distribution. Among other things, time series equations with volatility of cash flow drives the probability of default in credit analysis; it determines the value of real options; it is the reason for hedging risks through long-term contracts; and it is a key variable in determining value at risk. Volatility is generally expressed as a percentage. In rough terms, volatility can be defined as the standard deviation of the rate of return. The rate of return can be the rate of change in stock prices, oil prices, demand, cost or any other assumption in a financial model.

If the percent change in a variable comes from a normal distribution, volatility can be used to measure the probability that future values will be above or below a certain level. To see this, it is necessary to remember a little bit of statistics that most of us forget. In a normal distribution, the probability of being within one standard deviation above and below the mean is 68.27% and the probability of being within two standard deviations is 95.4%. If the mean is 4 and the standard deviation is 3, then there is a 68.27% chance that the observed value will be between 7 and 1. The standard deviation was 3 and the range is 4+3 and 4-3 – one standard deviation above and below the mean. The normal distribution is so convenient to use because any probability can be obtained if one knows only the mean and the standard deviation. Since volatility is the standard deviation of change in a variable, if one knows the volatility, the probably of achieving a value for the next year can be estimated. Say the volatility of oil prices is 20% per year (the actual volatility from historic levels before 2007) and the average oil price in January is $54 per barrel (the oil price at the beginning of 2007). Then there is a 68% chance that the oil price will be between $43.2 and $64.8 (an increase or decrease of 20%) by the end of the year. Similarly, there is a 95% chance that the oil price will be between $75.6 and $32.4 which are two standard deviations above and below the mean. These are the approximate actual values for 2007 oil

prices and volatility. By the way, the price at the end of 2007 was $96 per barrel, well outside of the 95% range.

Using Probability Distributions in Practice with Excel Rather than with Confusing Equations with Greek Letters

To understand time series equations, it is helpful to work through the mechanics of a couple of excel functions that compute the probabilities of normal distributions. In working with normal distributions, there are two functions that are useful in determining whether the rate of change in variables really comes from a normal distribution. The first function, the NORMDIST function, uses the mean, the standard deviation and an observation from the distribution as inputs and then returns probabilities. If the mean rate of change is 2%, the standard deviation is 15% and you would like to know the probability of realizing a rate of change of less than 20% you can use the cumulative option in the function as follows:

Cumulative Probability = NORMDIST(x = 20%, Mean of 2%, Standard Deviation of 15%, 1)

In the above equation, the switch of 1 at the end of the formula means that the probability output is a cumulative distribution and the X of 20% means that the output is the probability is less than or equal to 20% (NORMDIST(20%,2%,15%,1).) In this case the probability of achieving a growth rate of 20% or less is 88.5%. If a distance of one standard deviation above the mean or 17% is used (2% plus 17%), then the probability of being below this number is 84.13%, while if the observation is one standard deviation below the mean -- -13%, then the probability is 15.87%. The difference between these values found with the NORMDIST function (84.13% minus 15.87%) gives you the 68% chance of being within one standard deviation of the mean. It might be good to open a spreadsheet and try this out.

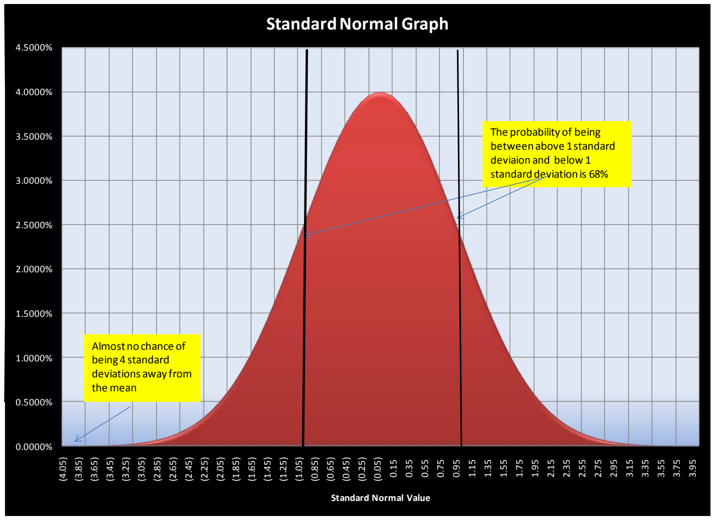
You can also use the NORMDIST function to compute the probability of falling into various increments on a non-cumulative basis. In the above example, you may want to find the probability of achieving 20% rather than the probability of the result being less than or equal to 20%. This can be accomplished by using the same NORMDIST function with a switch of FALSE (or zero) rather than TRUE. The problem with this formula when computing a non-cumulative number is that excel does not know the width of the range to use in computing the probability because the probability of having a very small range between .19999999 and .20000001 would be very small; but the probability of achieving a probability between 0 and 1.2 would be large. To resolve the problem, when using the NORMDIST function with you can simply multiply the formula by the increment desired and you will get the probability. In the above example, if you want the probability within 1% of the 20% number (i.e. between 19% and 20%), you would multiply the outcome of the NORMDIST function by 1%, yielding a probability value of 1.29%. If you were working with wind speed data that has a mean wind speed of 7 meters per second, a standard deviation of .5 meters per second, and you want the probability of achieving 8 meters per second, then the you would use the NORMDIST function for a range of 1.0 to get the value between 7.5 and 8.5. You can verify the non-cumulative method for applying the by creating increments of a variable and then inserting the probability into an adjacent column. As the whole non-cumulative business using the NORMDIST is a bit confusing, when you use the NORMDIST function it is a good idea to sum the probabilities and make sure the values add up to 1.0.

A second function that is useful when working through mechanical issues associated with stochastic modelling is a function in which you input the probability and you then find the number in the distribution associated with this probability. Using the numbers from the previous example, suppose you would like to find the percent change that has a 95% or a 99% probability, meaning that there is a 95% or a 99%

change that the actual value will be below the computed value. This is the essential concept of value at risk. In the above case with a 2% mean and 15% distribution, one can use the inverse of the normal distribution called the NORMINV function. The arguments for this function are the probability, the mean and the standard deviation (without a cumulative or non-cumulative switch). To find the percent change for which we can be sure that 99% of the time we will be below the number, one would enter NORMINV(99%,2%,15%). The resulting percent change is 37%. Using the oil price example, if the volatility is 20% and we want to be 99% sure that the price will be below a computed value, we would find the percent change in price is 47%, implying a 2007 year end price of $79.19 per barrel. Recall again that the actual price was $96. With a beginning price of the year price of $54, a volatility of 20% and an ending price of $96, one can use the NORMINV to show that the probability is 99.996%.

To see how the NORMINV function can be used, consider the example of wind power where confusing estimates are made for electricity production with a short-term P90 case and a long-term P90 case. These statistics measure the production or the capacity factor at which one can be 90% sure that the production will be exceeded. If the P50 case has a 25% capacity factor, the P90 case would have a lower value, say 18%. The short-term P90 case includes variation in power that occurs from changes in wind from year to year while the long-term P90 case averages out short-term fluctuations. The long-term P90 case includes factors such as badly estimating the effects of wind speeds at different heights (wind shear) and the effects of wind turbine at one turbine on other turbines (the wake effect). If the wind shear, the wake effect or a host of other variables are wrong, they will be wrong for the entire life of the project and there is no cyclical effect of high wind speeds in one year offsetting low wind speeds in another year. As the short-term P90 covers both the cyclical effects of wind speed changes and the estimation problems discussed above, there is a bigger dispersion in the short-term measure than the long-term measure. For example, if the P50 case is 25%, the short-term P90 case may have a value of 15% while the long-term P90 case may have a value of 18%. Using this data the standard deviation associated with long-term changes and the standard deviation associated with both long-term and short-term changes can be derived with the NORMINV function. As the P90 case has a 10% change of occurring, the NORMINV formula you can use a simple goal seek function to derive the standard deviation in capacity factor that will result in the targeted P90 cases. You can then use the standard deviation in Monte Carlo analysis discussed below.

Since any value of a normal distribution can be expressed as the mean and standard deviation, one can subtract the mean from each value of a distribution and divide the result by the standard deviation. This number gives you the number of standard deviations away from the mean for any value which is known as the standard normal distribution. In the earlier example, where the mean was 2% and the standard deviation was 15%, the standard normal value of 20% is (20%-2%)/15% or 1.2, meaning that the 20% value is 1.2 standard deviations from the mean. The probability of this value can be found with the NORMSDIST function (the S is included for the standard normal distribution), where one simply enters the standard normal value or the number of standard deviations away from the mean. A graph of the standard normal distribution created with excel is shown below (as usual, press the F11 key). To create this graph, simply enter the standard normal values beginning with -4 and increasing to +4 (the chance of a normal distribution falling outside of four standard deviations is very tiny.) Once the numbers are entered in a column, use the NORSMDIST graph so the graph (with a cumulative distribution).



Chapter 21: Taking Mystery out of Applying Time Series Analysis and Monte Carlo Simulation to Financial Models

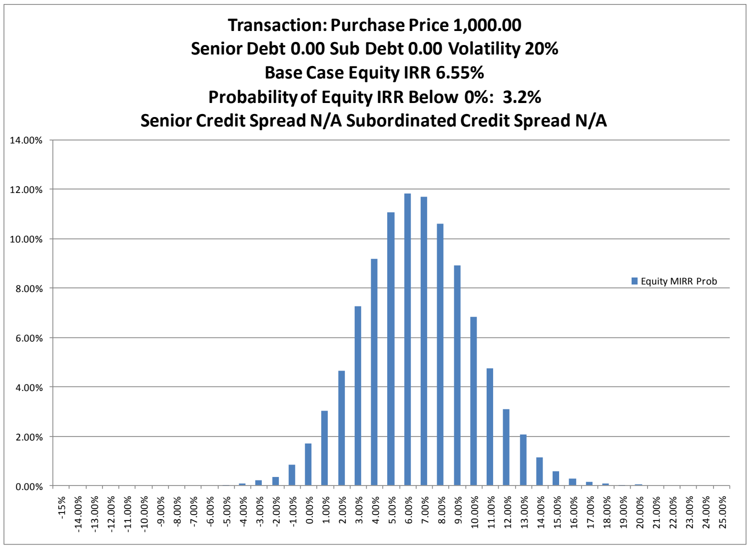
Some who have studied hard in university or who work with stochastic models and computer programs such as “At Risk” or “Crystal Ball” may develop a little arrogance the superiority of mathematical or stochastic models. Deterministic or non-stochastic models can be looked down on as analysis is somehow intellectually inferior. The idea of this and the next couple of chapters is to show how you can make elegant analysis with stochastic models and present statistics such as probability of loss, earnings at risk and many pictures of the distribution of financial ratios that would not be possible with deterministic models. Despite the beauty of the graphs, the attitude that creating a time series equation and running a simulation somehow produces better risk analysis is naïve and dangerous. The counter argument is that attempting to remove judgment from predicting how economic variables will move is simply impossible and analysis of underlying economic and business factors that drive key assumptions in a model cannot be avoided.

Before discussing the nuances of time series models, a step by step example of how one can easily create risk measures with Monte Carlo simulation is presented to remove any mystery of applying Monte Carlo simulation and time series analysis. The fact that mechanics of applying Monte Carlo simulation are not complicated does not mean that you should add Monte Carlo simulation to all of your financial models. Mechanics of Monte Carlo simulation are described so as to assure that you are not intimidated by the mechanics and the general idea of stochastic models and also to warn of the dangers simplistically applying statistical concepts to risk analysis.

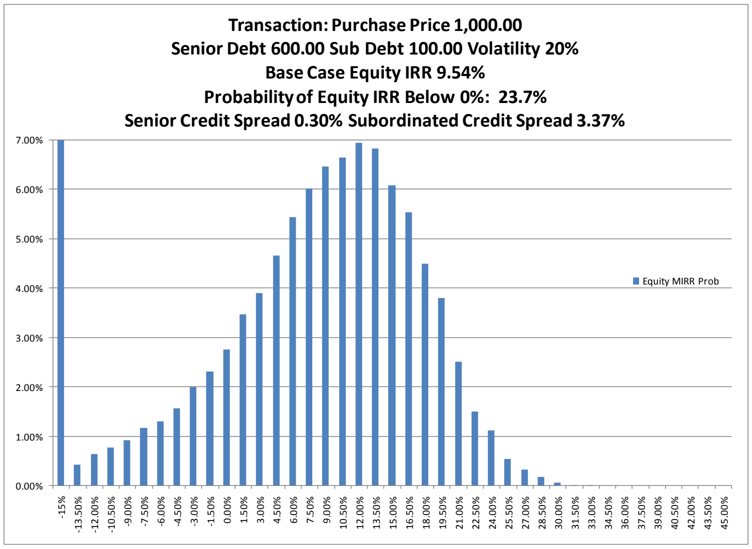
The step by step discussion below shows how you can integrate seemingly sophisticated risk statistics into financial models. The analysis presents the probability of earning more than the risk free rate and the required credit spread on alternative types of debt using different volatility parameters and different transaction structures. The only requirements for making this analysis are an equation that includes of volatility, knowledge of how to construct a simple time series model and a blank excel sheet (no special add-ins or VBA). The analysis uses a simple investment of 1,000 with operating cash flow of 150 and a

ten year life. Without financing and without volatility the modified IRR is 6.55% assuming a risk-free re-investment rate of 5%.

To add risk analysis to the investment, assume that the cash flows have a volatility of 20%. By including only the volatility factor, converting the 100 of cash flow into a time series equation and using Monte Carlo simulation, one can answer the questions like what is the probability that the equity IRR will be below zero. The graph below shows the distribution of IRR's that result when the volatility of cash flows is included in the analysis and demonstrates that it is possible to earn as much as 16% and that the probability of earning a return of less or equal to zero is 3.2%. Without debt, the distribution is not skewed.



To make the example a bit more interesting, assume the investment is financed with 600 of senior debt and 100 of subordinated debt. Without volatility, the base case IRR increases to 9.54%, but this statistic does not tell us about the risk to either debt holders or equity holders with debt in the financing structure. Through running the Monte Carlo simulation, one can see that the probability of realizing a return below zero increases to 24% and that there is a chance of earning as much as 27% rather than being limited to 16%. Further, the risk to the senior and subordinated debt holders can be measured. The senior debt needs a credit spread of .30% and the subordinated debt needs a credit spread of 3.37% to make the debt have equivalent returns to risk free debt. The large bar on the left of the graph measures the probability that the equity holders will loss their entire investment. Since the equity holders cannot lose more than their investment, the distribution of equity IRR is highly skewed, meaning that there is a limited downside and a wide variation of upside returns.



The results of the analysis change with different volatility parameters. If the volatility of cash flows is 30% instead of 20% then the upside and the downside for equity holders increase (the scale of the graph below is different from the prior graph.) The probability of a return below zero increases to almost 40%, but the relative chances of earning a very high return also increase. The increased volatility has a dramatic effect on the required credit spreads to debt holders. The senior credit spread increases from .24% to 2.02% and the minimum required spread on subordinated debt jumps to 15.39%.

To add this type of analysis to any financial model requires first reformulating a few of the equations from a financial model into time series equations and then evaluating the possible variability in selected output variables through using a random variable combined with a probability distribution and Monte Carlo simulation. The mechanical procedure for constructing the analysis is described in three sections below. The discussion begins with instructions on how to build a one column financial model used in the exercise that evaluates IRR's and the defaults on senior and subordinated debt. Next, the theory and mechanics of adding a simple time series equation into the financial model is addressed. The final section covers techniques which allow computation of risk statistics through making many different scenarios with different random numbers.

Step 1: Building a Flexible Deterministic Financial Model that Accepts Wide Variation in Cash Flow

The first and perhaps the most difficult step in the exercise of including Monte Carlo analysis is not the time series equation or the Monte Carlo simulation, but rather creating a financial model that computes the default on different types of debt and the rate of return earned by equity. The general method of creating a model with defaults using sub-totals and the MIN function was described in part one of the book in the context of the cash flow waterfall. Here the modelling ideas are converted to a one column model to illustrate Monte Carlo simulation through using the following four step process.

- First compute the future value of cash flows with no financing, showing the year by year cash flow in a single column over the 10 years;

- Second, compute the future value of the operating cash flows as well as the future value of the senior debt and the subordinated debt and place the values below the future value of the operating cash flows. For all of the future value calculations, a discount factor of 5% is assumed. The future value of the cash flows can be computed through using the SUMPRODUCT function and multiplying the cash flow by each future value factor. The future value of debt is the starting value of debt multiplied by the future value factor for 10 years;

- Third, compute the future value of the operating cash flows accruing to senior and subordinated debt. If the value of the future cash flows is less than the future value of the senior debt, then there is a default on the senior debt meaning that the value of the senior debt is the minimum of the nominal value of future debt or the future value of the cash flow. After computing the future value of senior debt, add a sub-total computing the value of the total cash flow less the value of the senior debt. Then, as with the senior debt, use the MIN function to compute the value of the subordinated debt, where the value is the minimum of the cash flow or the future value of the debt. These calculations could be computed using present value rather than future value. The only reason for computing future value is computation of the IRR discussed next.

- Fourth, compute the cash flow to equity as the future value of cash flow with no financing less the future value of senior and subordinated debt (this time no MIN function is necessary as it is the last line). The modified internal rate of return on equity can then be computed as the compound growth rate in the future value of equity cash flow relative to the equity investment (divide the future value by the current value and raise the product to one divided by the number of years). The only reason for using future value rather than present value in the above analysis is so that the IRR can be computed. As the future value is computed using a pre-determined re-investment of rate, the IRR is the modified IRR rather than the typical IRR computed with the IRR function.

An example of such a one column model is shown below. You can make a more elaborate model that displays separate rows for interest expense on senior debt and subordinated debt as well as the debt levels and the amount of defaulted debt. The reason for making a one column model is so that a Monte Carlo simulation can be created without any macros or VBA.

Year	FV Factor	Operating Cash Flow
0	1.63	(1,000)
1	1.55	100.00
2	1.48	100.00
3	1.41	100.00
4	1.34	100.00
5	1.28	100.00
6	1.22	100.00
7	1.16	100.00
8	1.10	100.00
9	1.05	100.00
10	1.00	100.00

FV of Operating Flows	1,257.79
FV Factor	1.63
FV of Senior Debt	1,140.23
FV of Sub Debt	325.78
Default on Senior	-
Default on Sub	208.22
Equity	-
Project MIRR	2.3%
Equity NPV	(95.24)
Project NPV	(216.98)
Equity MIRR	-100.00%

Step 2: Creating Time Series Equations for Key Assumptions in a Financial Model

To incorporate mathematical analysis such as probability of achieving a return below the risk free rate in the modelling process, the operating cash flow is modelled as a random variable where the potential variation is driven by the volatility parameter. The most basic way to create such a random is to assume the cash flow follows a random walk process that is a function of volatility and no other parameters. In a random walk, the variable (in this case cash flow) can move up or down from one period to the next with equal probability. One can think of this process by imagining a drunken man starting to walk along a line. After each step, the man may stumble in one direction or another with equal probability. After he takes the first stumble, the process begins again from the point of the last stumble, and the man can stumble again in each direction with equal probability. The prior stumble has no effect on the direction of subsequent stumbles. Depending on the how long many steps the man takes, the drunken man may wander in quite large lateral directions from the initial starting point. In this example, the range in size of each stumble can be though of as volatility. The random walk model can be described with the following equation of the current price purely a function of last period's price and random movement -- ε:

$$P_t = P_{t-1} + \varepsilon.$$

In this equation, ε term is a random term that can move up or down with equal probability and movements of ε in one period are independent of movements in other periods. This term can be replaced by the volatility percentage combined with a random number draw as shown in the formula below:

$$P_t = P_{t-1} + \text{Random Number} \times \text{Volatility Percentage} \times P_{t-1}.$$

When applying the simple equation above, the random number should have an expected value of zero (i.e. when a lot of random numbers are drawn, the average of the random numbers should average zero.) The current period price is a function of the prior period price and the random variable and volatility influences the price because it magnifies the effect of the random shock. To see how the equation works, if the volatility parameter is 20% and the random number happens to be 1.0, then the price would be 20% above the earlier price. If the volatility parameter were zero in the last part of the equation, there would be no movement in prices no mater what the random number. Finally, if the random number would be zero, then the price would not move no matter what is the volatility parameter.

To mechanically implement the time series equation, assume that the volatility is the standard deviation of the percent change derived from a normal distribution. In this case the number of possible deviations from zero should range from about -4 to +4 reflecting the potential number of standard deviations away from the average -- the most extreme cases are where the price moves by four times the standard deviation of 20%. To create the time series in excel with such a normal distribution, the functions RAND() and NORMSINV can be used together. The NORMSINV function accepts a number from zero to one and produces the standard normal value that reflects the possible number of deviations from the mean that are driven by the probability. Since the RAND function generates a number between zero and one, this function can be used to derive a random draw from a normal distribution.

$$\text{Cash Flow}_t = \text{Cash Flow}_{t-1} + \text{Volatility Parameter x (NORMSINV(RAND()) x Cash Flow}_{t-1}$$

When the equation above is plugged into excel, if the random number is .5, then the result of the NORMSINV is zero and the whole term drops out. Only in very extreme probability cases very close to zero or very close to 1.0 does the NORMSINV produce values anywhere near +4 to -4. Due to the characteristics of a normal distribution discussed above, there is a 68% chance that the NORMSINV(RAND()) factor will produce a value between +1 and -1.

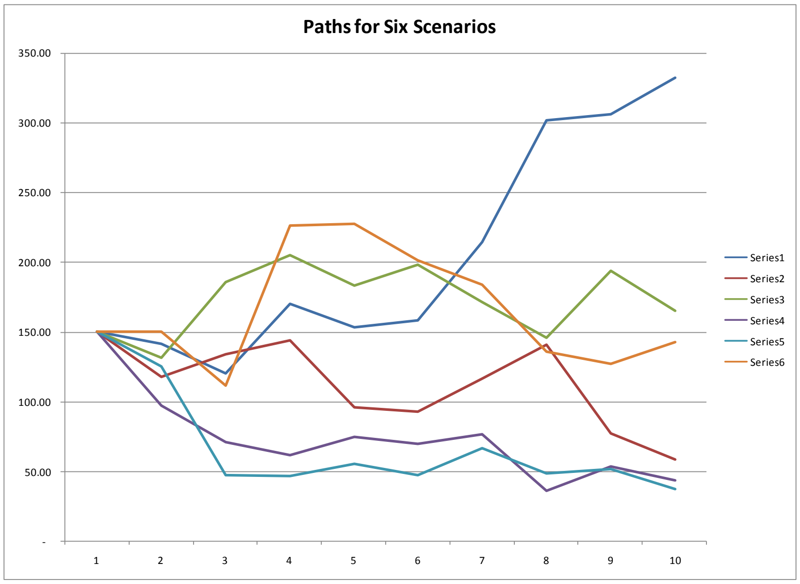
To implement this formula in the one column financial model, simply change the formula in the second year (i.e. the second 100 in the above diagram.) Instead of making the second cash flow equal to the first cash flow, add a term that allows the cash flow to vary depending on a random draw. For example, if the first cash flow is in cell C5, the formula for the next cash flow in C6 is C5 + C5 x volatility x NORMSINV(RAND()). Once this formula is entered for the second year cash flow, copy the same formula to all of the remaining rows meaning a new random number will be drawn for each year after the first year.

Step 3: Adding Monte Carlo Analysis with VBA Code

Once a time series equation is defined and a probability distribution such as the normal distribution is selected, the financial model can be used to create probability distributions of model outputs through using Monte Carlo simulation. Monte Carlo simulation was supposedly originally used in development of the Atomic Bomb in the 1940's (although without computers somebody would have had to spend a lot of time making up random numbers). The general idea of Monte Carlo simulation is to construct a large number of possible scenarios by using multiple random numbers together with a time series model where the movements in a variable are driven by standard deviation or probability.

Using the time series model above, Monte Carlo simulation can be created by copying the same column over and over again creating many different scenarios with different random numbers. In excel 2007 and later versions, there are more than 16,000 columns meaning that using the SHIFT, CNTL, → and CNTL, R technique to copy the model across the page (described in part 1) results in more than 16,000

different scenarios, each scenario having nine different random numbers (with this many scenarios, excel becomes very slow). The first six cash flow scenarios from this copying process (assuming 30% volatility) are shown on the chart below. Note how some scenarios result in a cash flow near zero while others result in very high cash flows. In general, the variance in prices for different scenarios increases as time passes (which should be the case for a random walk process).



Instead of using the SHIFT, CNTL, → functions to create many different scenarios, you can also write a little macro that repeats the process over and over again in a similar manner. One of the advantages of using such a macro is that you can keep the output from the simulation in a separate sheet and you do not have to repeat the whole financial model as in the one column example (meaning that you do not have to create a one column model). To create this, combine the FOR NEXT loop with the CELLS statement in a similar manner as discussed in the break-even analysis chapter above for creating a data table with VBA. The following code will perform the simulation and place outputs in a sheet named output_sheet beginning in column 5 using these statements:

```
num_of_simulations = INPUTBOX(" Enter Number of Simulations")
For Row = 10 to num_of_simulations

        Sheets("output_sheet").cells(Row, 5) = Range("Project_IRR")
        Sheets("output_sheet").cells(Row, 6) = Range("Senior_default")

        Further rows for recording other variables ….

       Application.StatusBar = " Iteration Number " & Row
Next Row

Application.StatusBar = FALSE
```

Once the results of multiple scenarios are placed somewhere in the workbook, the PV, the IRR, and the defaults can be summarized. Some of the scenarios result in a relatively high IRR and some a low IRR; some result in no default and some result in a large default. The average level of default can be divided by the level of the debt to compute the required credit spread on senior debt and subordinated debt. The frequency graphs illustrated above can be computed using FREQUENCY function as follows:

- Compute the maximum and minimum values across all scenarios for the variable that you would like to graph or compute the mean plus or minus four standard deviations from the raw data computed from copying the column or the macro. For example, if you are graphing the NPV, compute the maximum and minimum NPV.

- Create a data list of bins that will accumulate the number of scenarios in the various increments and list the increments in a row or a column. If the minimum is -200 and the maximum is 1,000, then you could make increments beginning with -200 and increasing by 50 until 1,000 is reached.

- Shade the row or the column next to the increments and enter the FREQUENCY function which accepts the range in values and the bins. When you are finished entering the function press the SHIFT, CNTL, ENTER sequence.

- Once the bins and the frequency are entered, you can use the F11 key to make a graph and/or compute the probabilities through dividing the frequency by the sum of all of the observations. If there are no other numbers in the row ore the column of the frequency data, you can use the sum function for the entire row or column as illustrated below:

Probability = C10/SUM(C:C)

The process of using the FREQUENY function to create probability distributions is illustrated below. Selection of the number of bins determines the increment and depends on the number of scenarios that are created with the Monte Carlo simulation.

Pre Tax IRR

	Pre Tax IRR	Pre Tax MIRR	NPV at 5%	NPV at 2%	PV of Debt Default	Debt Default Percent	Equity MIRR	Pre Tax IRR
Average	16.79%	10.44%	34,595,618	54,061,574	110,180	0.66%	12.14%	0.17
Standard Deviation	8.74%	2.48%	25,830,583	36,059,706	749,809	4.46%	16.21%	0.09
Count	1001	1001	1001	1001	1001	1001	1001	1,001.00
Incements	20	20	20	20	20	20	20	20.00
Std Devs	4	4	4	4	4	4	4	4.00
Lower	-18.16%	0.54%	(68,726,714)	(90,177,251)	(2,889,055)	-17.20%	-52.70%	(0.18)
Upper	51.75%	20.35%	137,917,949	198,300,398	3,109,416	18.51%	76.97%	0.52
Min	-45.96%	2.59%	(15,819,851)	(16,789,340)	-	0.00%	-100.00%	(0.46)
Max	36.81%	15.68%	113,026,479	162,037,128	10,540,052	62.74%	21.90%	0.37
Start	-18.16%	2.59%	(15,819,851)	(16,789,340)	-	0.00%	-52.70%	(0.18)
End	36.81%	15.68%	113,026,479	162,037,128	3,109,416	18.51%	21.90%	0.37
Difference	54.97%	13.09%	128,846,330	178,826,469	3,109,416	18.51%	74.60%	0.55
increment	2.75%	0.65%	6,442,317	8,941,323	155,471	0.93%	3.73%	0.03

Increment	Probability	Frequeny	Adjusted
-18.16%	0.50%	5	5
-15.41%	0.00%	0	0
-12.66%	0.00%	0	0
-9.91%	0.10%	1	1
-7.17%	0.10%	1	1
-4.42%	0.30%	3	3
-1.67%	0.40%	4	4
1.08%	0.90%	9	9
3.83%	1.80%	18	18
6.58%	3.40%	34	34
9.33%	7.69%	77	77
12.07%	10.69%	107	107
14.82%	12.49%	125	125
17.57%	14.99%	150	150
20.32%	13.99%	140	140
23.07%	10.39%	104	104
25.82%	9.49%	95	95
28.56%	5.79%	58	58
31.31%	4.20%	42	42
34.06%	1.90%	19	19
36.81%	0.90%	9	9
39.56%	0.00%	#N/A	0
42.31%	0.00%	#N/A	0
45.05%	0.00%	#N/A	0
47.80%	0.00%	#N/A	0
50.55%	0.00%	#N/A	0
53.30%	0.00%	#N/A	0
56.05%	0.00%	#N/A	0
58.80%	0.00%	#N/A	0
61.54%	0.00%	#N/A	0

Chapter 22: Constructing Probability Distributions and Demonstrating the Dramatic Importance of Including Trends, Mean Reversion, Price Boundaries, and Correlations among Variables as well as the Volatility Parameter in Time Series Equations

This chapter turns to more complex aspects of time series models that include only parameters for time trends, mean reversion, lower and upper boundaries, jump processes and correlations. The time series equation used in the last chapter with only a volatility parameter would generally be inappropriate for use in financial models because it assumes a normal distribution of random numbers; no mean reversion, no upper or lower price boundaries and no interdependence between different variables. Actual Monte Carlo simulations should be more complex than the simple example with only volatility because the time series equations include added parameters (trends, boundaries, jumps and so forth), because of transformation to logs, and because the random draws may be filtered through a probability distribution other than the normal distribution. In the more complex case, the time series equation (without jumps, price boundaries or correlations and not in logs) can be written as:

$Price_t = Price_{t-1}$ x Trend Factor + ($Price_{t-1}$ - Average Price adjusted for Trend) x Mean Reversion Factor + Volatility Parameter x Inverse Normal Distribution (Random Draw) x $Price_{t-1}$ + Adjustment for Correlation

While this equation is more complex than the simple equation, no matter how the time series equation is specified and no matter what probability distribution is used, the process always boils down to the idea of drawing a random number, adjusting it for volatility and making the next period price a function of the prior period price. For purposes of the discussion in this chapter, parameters such as volatility, mean reversion factors, correlations and other factors are assumed to be given.

The Starting Point for Developing Time Series Equations -- Brownian Motion and Normal Distributions with are Not Representative of Most Economic Variables

The most common set of time series models applied to financial instruments such as stocks, options and other derivatives are random walk models built on assumption that rates of change follow a normal distribution and that changes in price are independent of past prices (for example in the Black-Scholes model an underlying assumption is that prices follow a normal distribution and historic price changes have nothing to do with prospective price changes.) A fundamental property of these models is that prices can be called non-stationary, meaning that they can wander all about without coming back to some average level.

Efficient markets theory is consistent with the notion that stock prices follow a random walk process (price changes should be independent with one another.) The general idea of efficient markets theory is that any change in stock prices arises from new information – not earlier information which has already been incorporated in the price. This implies that future price changes have nothing to do with price changes that occurred in earlier periods. If new information is normally distributed, and prior information has already been included in prices, the price process follows a random walk process and where past prices -- P_{t-2}, P_{t-3}, and so forth have no forecasting value. If past prices are irrelevant, there is no mean reversion in prices; prices are only correlated with other prices that have also have random walks and in terms of the rate of return, there is no upper or lower boundary. While this theory may apply to stock

prices, it clearly does not apply to many variables that are entered into financial models and that often drive risk assessment. For example, in the case of electricity prices, a change of one direction in the price of electricity often means that subsequent prices will at some point in the future change in the opposite direction. In this case past electricity prices are clearly very relevant in predicting future price changes suggesting that electricity prices do not follow a random walk.

The common equation used in financial economics is similar to the equation used in the simple example in the last section in which random draws are extracted from a normal distribution. However, there is slight technical difference involving transforming the data to logarithms. This difference involves the assumption of the rate of changes in prices following a continual process rather than happening at discrete time periods. With continual growth, the price in the next period is not defined as this period's price multiplied by one plus the volatility as in the above equation, but rather this price is the exponent of volatility as shown below.

Discrete:	$P_t = P_{t-1}$ x (1+ P_{t-1} x Volatility Percent x Standard Normal Draw)
Continuous:	$P_t = P_{t-1}$ x (exp(Volatility Percent x Standard Normal Draw))

If the time increments are continuous rather than discrete, the random walk becomes Brownian motion. In Brownian motion, the ε term is a draw from a normal distribution with a variance that increases on a linear basis over time. Use of the continuous distribution prevents the possibility of a negative price (the exp function cannot be negative) and involves somewhat more complicated mathematical equations. Rather than burdening you with these more complex formulas which do not really effect the most of the fundamental risk analysis issues, the continual distribution is discussed in Appendix A at the end of the chapter.

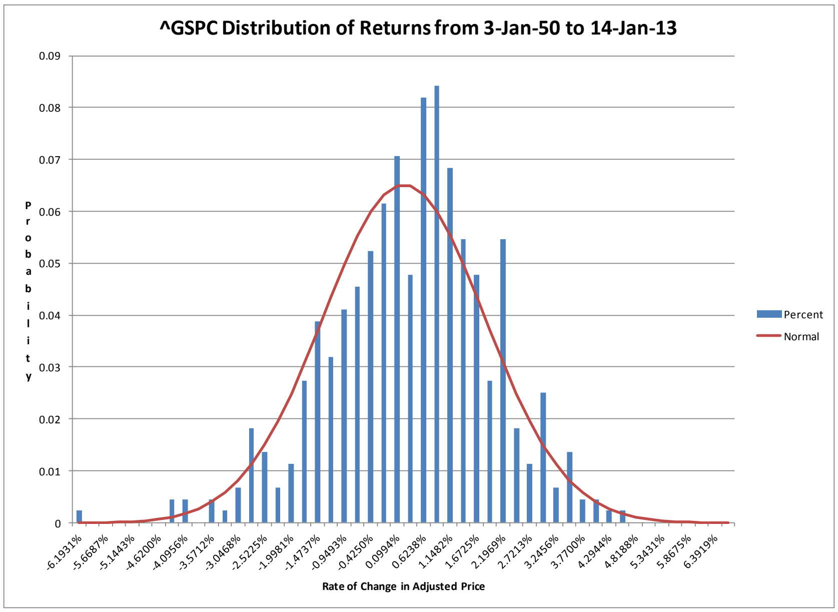
Attempting to Test the Basic Assumption that Rates of Change in a Variable Come From a Normal Distribution

One of the basic questions before deciding to implement a model with Brownian motion is determining whether the underlying assumption of a normal distribution and independence of price changes is applicable. In testing whether rates of change follow a normal distribution, the NORMDIST can help you understand whether a distribution is normal and thus whether it is appropriate to apply draws from a standard normal distribution in a Monte Carlo analysis. To compare the distribution of a particular variable to a normal distribution, the following steps can be followed:

- Compute the percent change in the series and the mean and the standard deviation of the series of percent changes
- Create bins for a frequency distribution. You can begin with something like the mean minus the standard deviation times 4 which would be a very extreme value for the normal distribution. Then, you can enter the number of increments and divide the standard deviation multiplied by 8 to compute the increment.
- Compute a frequency distribution and the probability of the series of percent changes (divide the frequency by the total number of observations to convert the frequency distribution into a probability distribution. To do this, shade the area next to the bins, enter the FREQUENY distribution and press SHIFT, CNTL, ENTER
- Compute the probability distribution for a normal distribution using the NORMDIST function with a switch for a non-cumulative distribution. Then multiply the formula by the increment.

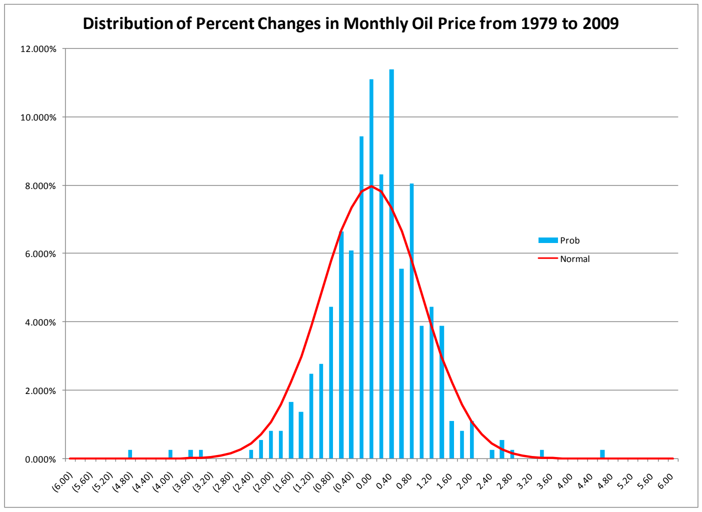
- Use the F11 key to make a graph of the percent changes relative to the normal distribution.

Once you get used to this process, it will go very quickly and you can test all kinds of things. For example, if you want to test the distribution of wind over a year, you compute the mean and the standard deviation and then you can enter a normal distribution in the same manner. The results for daily changes in the value of the S&P 500 are shown in the graph below. Without making any highly sophisticated econometric analysis, the graph below illustrates that daily stock price changes have not been normal.



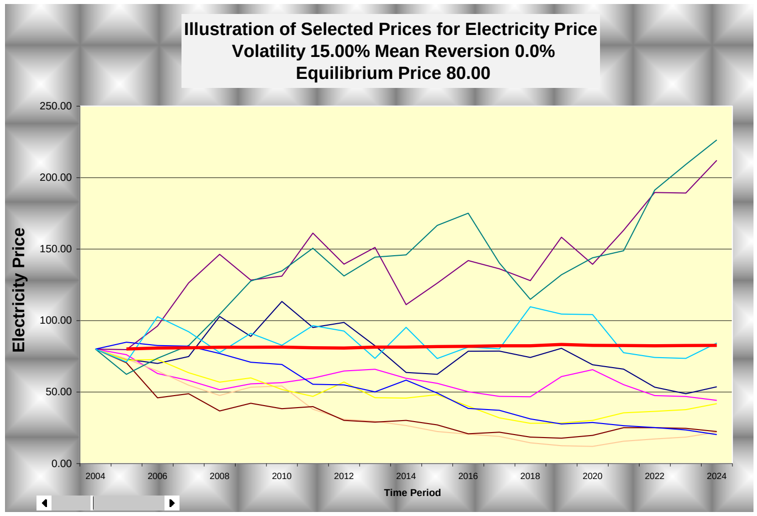
Since changes in stock price do not follow a normal distribution, one could attempt to fit another distribution, but that would be difficult and dangerous. The danger comes from the very small chance that the percent change is price could be six times away from the mean. Since 1950, there have been 17 days that have been in the negative six bucket. In a normal distribution, the chance of reaching minus six standard deviations away from the mean is one day in 4.02 million years (this can be computed using the NORMSDIST function, plugging in -6 and then multiplying the probability by 252 days per year).

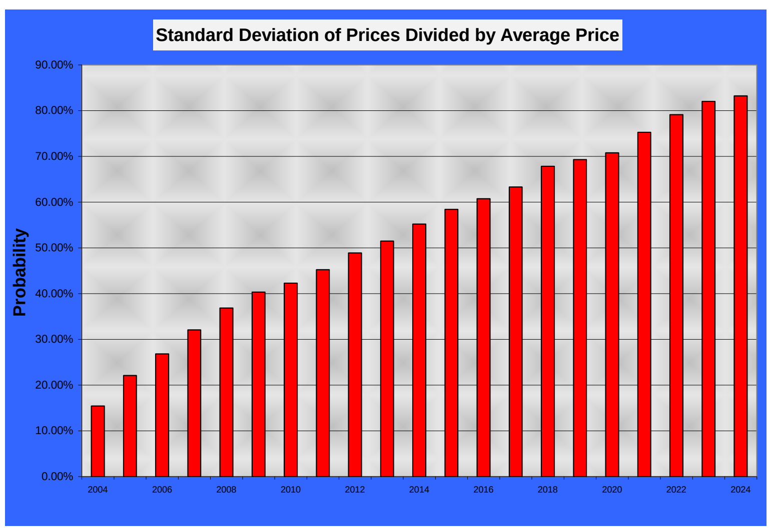
The story is similar for the distribution of oil prices. The graph below shows the distribution of the monthly rate of change in real oil prices compared to a normal distribution. As with the S&P 500, there are a number of observations that exceed four standard deviations and would be almost impossible if monthly oil price changes were really normally distributed.



Results of Unrealistic Brownian Motion Case with No Price Boundaries, Mean Reversion or Correlations

To demonstrate the effects of mean reversion, price boundaries, correlations and other factors a case study of a commercial real estate building is used. This case is used to evaluate different strategies with respect to signing long-term leases. In this case Monte Carlo simulation can be used to evaluate the trade-off between realizing shorter leases with higher rental rates versus longer leases with lower rentals. The case begins with a simple simulation that only includes the volatility parameter which is unrealistic for most variables (even arguably stock prices). This case is presented as a baseline to illustrate the dramatic effects of adding the other parameters. In introducing the various factors, the patterns of prices resulting from the simulation are introduced followed by a summary table describing the financial results. The first graph below shows results of a few of the price simulations in the case with volatility of 15% while the second graph shows graph divides the standard deviation of prices in the various simulations by the average price across the simulations. Even though the volatility is only 15%, the first graph shows that the price can eventually move from 80 to 200 while in the low cases the price is near zero. As expected, the variation of prices increases through time. When all of the simulations are averaged for each period, the price is just about equal to the beginning price as expected. Finally, with Brownian motion, the standard deviation divided by the initial price increases as the length of the period increases.





Financial results from the Brownian motion case are shown below. The probability of loss on the debt is 45% and the expected loss is 23% of the debt. The chance that the equity IRR will be below 5% is 31% (measured by computing the NPV at 5% and counting the negative NPV values). Because of the many scenarios with very low prices, the average IRR is negative and the standard deviation is very high relative to the average NPV.

Time Series Parameters	
Starting Price	7.00
Volatility	7.00%
Minimum Rate	0.00
Maximum Price	0.00
Mean Reversion	0.00%
Mean Price	8.00

Model Outputs	
Equity IRR	-6.37%
Pre-tax IRR	1.96%
Probability of Default	45.75%
Expected Loss as Percent of Debt	23.48%
Probability of IRR < 5%	31.57%
Probability of IRR < 2%	28.27%
Std Deviation as Pct of Avg NPV	192.59%

The Dramatic Effects of Price Boundaries and the Idea that Prices Generally Do Not Fall Below Short-run Marginal Cost for Sustained Periods of Time

In a time series equation, the price can suddenly or gradually fall to low levels or reach very high levels. Because of the mathematics of the time series equation, after the price reaches a low level, the volatility is applied to the low level and prices tend to remain at low levels. The opposite tends to occur with high prices where the volatility parameter is magnified after high prices are reached. If prices would really reach the extreme values in actual markets, the prices would prompt responses by consumers or suppliers that effectively put limits on the prices. In the case of low prices, limits occur because of producers ceasing production. In the case of high prices, limits come from consumers who reduce demand. If prices fall below short run marginal costs, companies will eventually chose not to produce. The reduced supply increases price and limits the decline in price to short-run marginal cost. This implies that prices in time series models should have a lower boundary defined by short-run marginal cost. Upper price boundaries can also be appropriate in time series models because if prices reach high levels, demand may be curtailed by factors such as substitute products.

Boundary conditions can be incorporated in time series models through placing limits on the lower or upper possible prices. The models can simply include a conditional factor that when prices drift above the upper boundary or below the lower boundary, the price is set to the boundary level. As with the other components of a time series model, the boundaries themselves could be stochastic random variables with volatility and mean reversion. Implementing lower and upper boundaries can be accomplished through using a MIN function and a MAX function.

To illustrate the effect of price boundaries, the real estate case is adjusted for a lower bound on rental rates and an upper bound that is very low relative to experiences in the market. The effect of boundaries on measures of financial risk dramatic as illustrated on the table below. Even though the minimum price is a very low level of 3.0, the probability of loss on the debt falls to 2.1%. The chance that the equity IRR will be below 5% is 10.8% instead of 31%. Because of there are less scenarios with very low prices, the average IRR is 10.77% and the standard deviation is much lower relative to the average NPV.

Time Series Parameters	
Starting Price	7.00
Volatility	7.00%
Minimum Rate	3.00
Maximum Price	15.00
Mean Reversion	0.00%
Mean Price	8.00

Model Outputs	
Equity IRR	10.77%
Pre-tax IRR	16.28%
Probability of Default	2.10%
Expected Loss as Percent of Debt	0.24%
Probability of IRR < 5%	10.79%
Probability of IRR < 2%	3.30%
Std Deviation as Pct of Avg NPV	94.66%

The General Idea of Mean Reversion and Long-Run Equilibrium Analysis and why Mean Reversion should be Present in Many Variables used in Cash Flow Models

When an economic variable exhibits cyclical behaviour or reverts to average levels, the variable can be modelled as a mean reverting process rather than as a random walk process. For time series equations with mean reversion, variables still have random shocks driven by volatility; but after a random shock causes the process to move away from mean levels, prices in subsequent periods tend to move back to a defined average or equilibrium price. The speed at which prices revert to the long-run equilibrium after a random shock can be expressed in terms of an annual percentage. If the long-term equilibrium or average price is 100 and a shock causes prices to move to 150, then a mean reversion factor of 40% would mean that the next year price tends to move down by 50 x 40% or 20. The basic equation for a mean reverting process is:

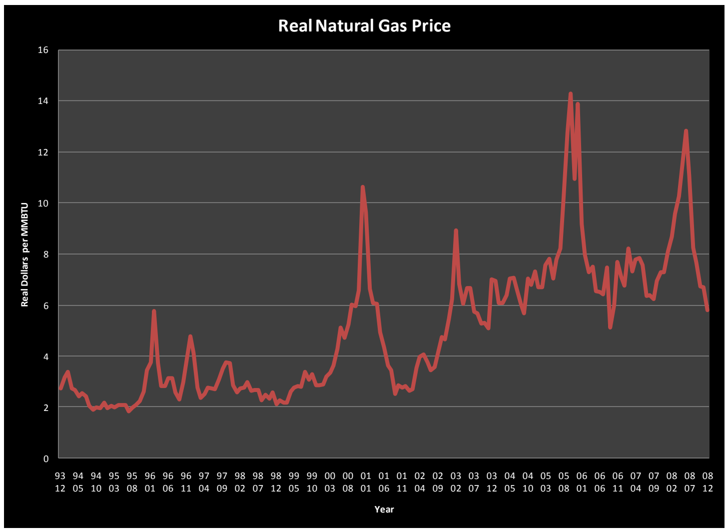
$$P_t = P_{t-1} + \text{Mean Reversion Factor} \times (P_m - P_{t-1}) + \varepsilon$$

In a time series formula with mean reversion, the mean price term P_m should in theory approximate the long-run cost of production and the ε term incorporates volatility and draws from a normal distribution as was the case for random walk processes using the NORMSINV(RAND()) functions. The mean reversion factor in the equation should generally between zero and one (although it could theoretically be negative). If the mean reversion factor is 1.0, then the P_{t-1} terms cancel and the equation becomes $P_t = P_m + \varepsilon$ implying that prior period observations have no influence on the forecast; everything starts again from the average. In this case, there are movements away from the mean level due to random shocks, but the process starts over at the mean level in the next period – one could think of variations in wind and solar power as following a process like this. If the mean reversion factor is zero, the equation is the same as the random walk processes described above where prior period prices have no forecasting value and the process is non-stationary.

If the prices are expressed in annual terms, one divided by the mean reversion factor in the equation can be thought of as the amount number of years it takes to come back to the equilibrium level. With a mean reversion factor of .5, it takes two years on average to revert to the mean. Therefore, one can think of the mean reversion factor in terms of how many years it takes for a market to reach equilibrium after a random shock. For example, if a market has deficient supply resulting in high prices and it takes three years to construct new plants, then the annual mean reversion factor should theoretically be around 33%. Alternatively if there is excess supply and demand growth would cause take ten years until a supply/demand balance occurred, then the mean reversion factor would be one divided by ten or 10%.

Variables which move according to Brownian motion wander around and gradually move more and more away from their initial value. While this behaviour may be characteristic of some variables such as stock prices, many if not most economic variables move in cycles rather than in random walks. In the case of price variables, the cyclical movement back to an average value is due to the simple fact that high prices will prompt increased supply thereby moderating the price increase while very low prices will cause new supply to slow down and demand to increase. In the case of electricity where storage is not generally available, the mean reversion comes from fact that demand itself is mean reverting due to the weather and due to the steep slope of the supply curve. Most commodity prices including oil, gas, real estate, iron, copper, coal and electricity should eventually move in the direction of their long-run cost of

production rather than continually moving up or down without limits. This tendency of oil prices is noted by Dixit and Pindyck as follows: "…while in the short-run the price of oil might fluctuate randomly up and down, in the long-run it ought to be drawn back towards the marginal cost of producing oil."[3] The propensity to move to equilibrium amounts is known as mean reversion. The tendency for natural gas prices to gradually revert to average levels is illustrated in the graph below where prices seem to increase for a couple of years and then fall back.



For electricity prices, mean reversion is a particularly prominent statistical feature over both short-term and long-term periods. The mean reversion of electricity contrasts starkly with stock prices which, if you still believe in efficient markets, supposedly have the property that historic prices do not influence future price changes. Market capitalisation of a company is a measure of the store of the value of a company created by cumulative past decisions and past events as well as expected future cash flow from current management strategies. In theory, stock prices change only when new information arises that was not already reflected in the current price. Electricity, on the other hand, cannot be stored meaning that once an event occurs that affects prices such as hot weather or a plant outage, the next period price starts over at production costs with normal weather or plants back in service. Electricity prices are the polar opposite of stock prices in terms of the relevance of past prices and the tendency to revert to mean levels.

The reason mean reversion is often present in commodity prices is because of the way new capacity is developed when prices move. In many industries including real estate, telecommunication, infrastructure and commodities, the behaviour of developers is somewhat unpredictable and it is arguably irrational in the short-term. However, after surplus capacity has been constructed, at some point construction will slow or because of low prices. This behaviour ultimately forces prices to move in the direction of long-run marginal cost. The behaviour of developers can be modelled as a random process that gradually moves back to equilibrium levels driven by production costs, but does not remain at stable levels. The fact that new developers of a project cannot indefinitely lose money (because no

[3] Ibid, Dixtit and Pindyck, page 74.

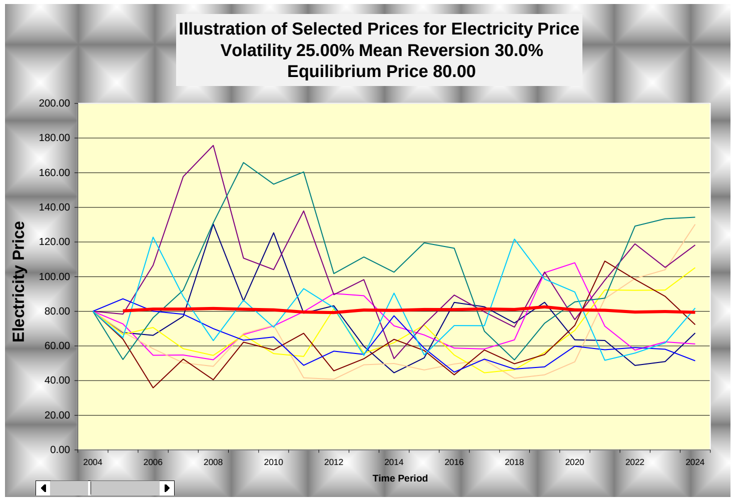
new investments would occur and reduced supply would push prices up) or continually earn supernormal economic profits (because of new entrants would be attracted to the market pushing prices down) is represented by movement of prices to the long-run marginal cost. To create a mathematical model of economic behaviour that pushes prices move to long-run marginal cost, a mean reversion factor can be included in time series analysis.

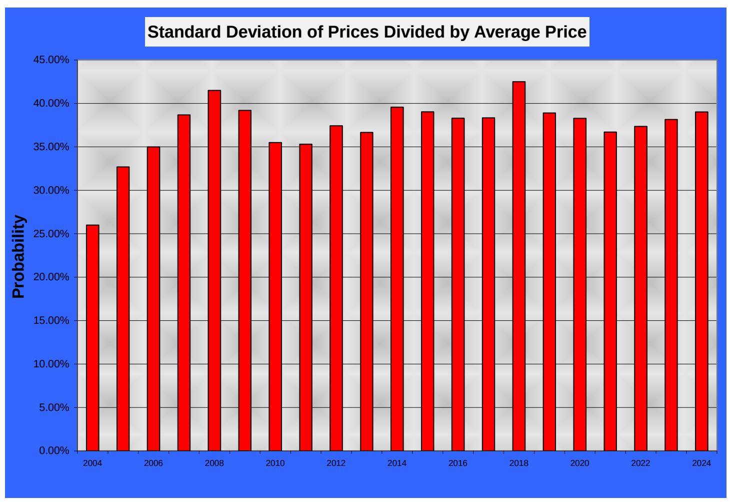
One further complication is that the equilibrium price or long run marginal cost may be random itself. In the case of electricity prices, the long-run marginal cost is often highly influenced natural gas prices which have high volatility. To implement a model where equilibrium prices follow a random process, a two factor model can be developed. Here, two random draws are made and a volatility parameter applicable to the equilibrium price is included in the time series model. The long-term marginal cost or equilibrium price that is input in a time series model can also vary from period to period.

The Dramatic Effect of Mean Reversion Parameters on Risk Distributions when Running a Monte Carlo Simulation

When applying time series analysis and Month Carlo simulation to investments, mean reversion parameters can have a dramatic effect on the various risk measures. To illustrate this, consider the real estate example introduced above. In addition to the volatility parameter, a mean reversion parameter and a mean price is included in the simulation. As in example introduced in the last chapter and in the Brownian motion case, the first step is constructing the time series equation, the second step is running a Monte Carlo simulation to obtain scenarios with different prices and the final step is using the different prices in the financial model to obtain a distribution of IRR's, minimum credit spreads or other financial variables.

Simulations of prices that include a mean reversion parameter demonstrate differences in the characteristics of a mean reverting time series from a random walk series. The exercise of simulating prices with mean reversion allows one to inspect the simulation results and evaluate whether the price patterns reasonably represent possible movements in real estate prices. Outcomes of the case with mean reversion of 30% are presented on the two graphs below. The first graph illustrates that with mean reversion, the extreme price cases are far more unlikely. Unlike the case of a random walk process, the standard deviation of prices in a series with mean reversion does not increase over time.





Adding mean reversion factors dramatically changes the measures of risk relative to the cases without mean reversion. Even if an annual parameter of only 10% is assumed -- which assumes that it takes 10 years to reach equilibrium through changes in construction of new plant or demand increases -- risk is reduced dramatically. Using the same real estate as above and removing the lower and upper bounds the small mean reversion parameter has dramatic effects on the credit statistics and the variation in return on equity. Without mean reversion, the probability of default falls from 45.7% to 8%. In the case with only 10% mean reversion, the probability of achieving a return below 5% decreases from 32% down to 5%. If mean reversion is 33% implying three years to reach equilibrium, the risk for senior lenders or subordinated lenders is to almost zero. Clearly the presence of mean reversion has a dramatic effect on the risk analysis and coming up with an appropriate parameter is essential. It should be clear from the table below that ignoring mean reversion or over-estimating mean reversion can lead to big mistakes in evaluation of the risk of an investment.

Time Series Parameters	
Starting Price	7.00
Volatility	7.00%
Minimum Rate	0.00
Maximum Price	15.00
Mean Reversion	0.83%
Mean Price	7.00

Model Outputs	
Equity IRR	12.14%
Pre-tax IRR	16.79%
Probability of Default	7.99%
Expected Loss as Percent of Debt	0.66%
Probability of IRR < 5%	5.29%
Probability of IRR < 2%	2.50%
Std Deviation as Pct of Avg NPV	74.66%

Modelling Correlations among Variables in Time Series Equations

Up to this point, the time series processes have been discussed for one variable without regard to how the movement is affected by changes in other variables. In most situations, analysing a time series equation in isolation is inappropriate because key variables are correlated with another and a movement of one variable will affect the way other variables change. For example, in the case of the electricity plant fired with natural gas used to demonstrate the effect of alternative time series equations there would likely be a correlation between natural gas prices and electricity prices. If a time series equation is used in forecasting the price of electricity, the price of natural gas should have related movements in another time series model because of the ability to substitute fuels, because natural gas plants are often marginally running plants that drive the price and because the price of natural gas influence the type of new plants that are constructed.

The relation between natural gas and electricity prices is so important that the difference between electricity price and natural gas price is often modelled and tracked separately from either the electricity price or the gas price. This difference in price is named the spark spread as shown on the formula below.

Spark Spread = Electricity Price ($/MWH) – Gas Price ($/MMBTU) x Heat Rate

Using the spark spread to model the cash flow of a gas plant presumes that there is perfect correlation between gas and electricity prices. In some markets, there is very high correlation between natural gas and electricity prices but in other markets the correlation is relatively weak.

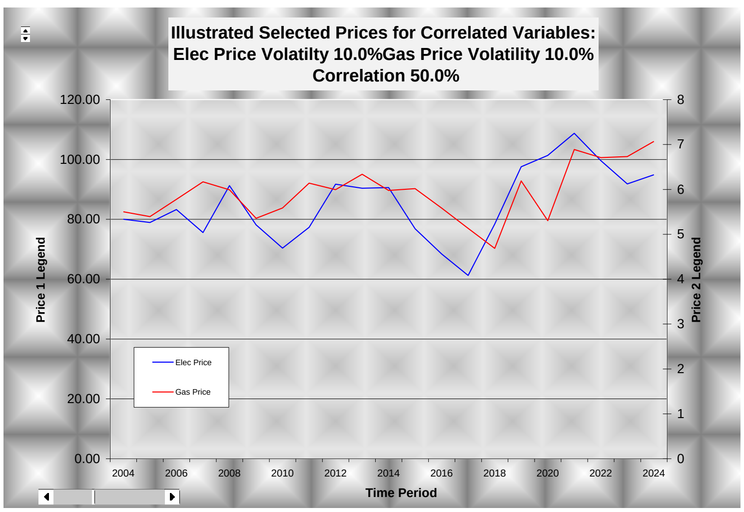
In the example above where there is lack of perfect correlation between natural gas prices and electricity prices, an alternative way to analyse distributions of cash flow for the gas plant is to create a time series models of both gas and electricity and to incorporate the correlation between the prices in the model. The process of incorporating the interrelationship between variables in time series models begins by defining a parameter that specifies the correlation between variables. The linear correlation coefficient has a value between 1 and -1 and is defined by the following formula:

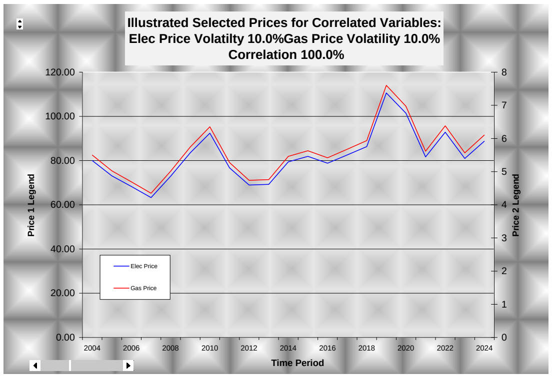
$$\text{Correlation}_{1,2} = [\text{Covariance}_{1,2}]/[(\text{Variance}_1 \times \text{Variance}_2)^{1/2}] = \rho_{1,2}$$

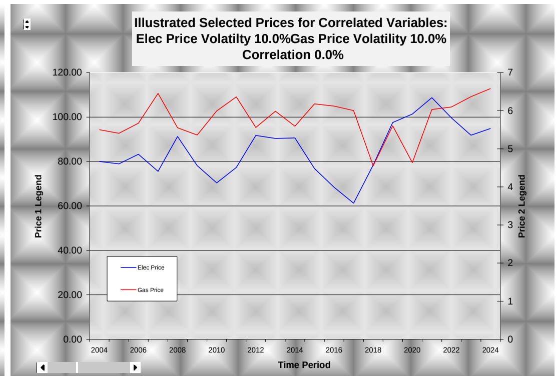
The correlation coefficient combined with volatility parameters can be used make a forecast of how the natural gas price and the electricity prices move together. The procedure for incorporating correlation into a time series model involves making random draw for the second correlated variable be a function of the random draw for the first variable. The mechanical procedure for making these calculations using

something called the Cholesky function is somewhat dense and therefore relegated to Appendix B. This appendix attempts to explain in relatively simple terms the procedure for computing the random variation in one variable that is driven by the variation in another.

To illustrate the effect of reflecting correlation in variables in the analysis, the natural gas price in the earlier example is modelled as a stochastic variable rather than as a fixed amount as in the previous cases. Once reflecting volatility and mean reversion in the natural gas price, various different correlation factors are used to illustrate the effect of including correlation in the time series analysis. The effect of assuming no correlation, 50% correlation, 100% is evaluated in analysis illustrated below. For illustrating the effects of correlation, the case with 40% volatility and 10% mean reversion is used for both electricity and natural gas prices. In the first example, the natural gas prices are assumed to not be correlated with electricity prices, while in the second example, the correlation is assumed to be 50% and the in the third case the correlation is 100%. Selected price paths of the electricity prices and natural gas prices assuming different correlations (with volatility of 10%) are shown below. Note that in the case with 100% correlation the prices closely move together while in the case with 0% correlation there is no relation and in the case with 50% correlation, the prices generally move together, but the not all of the moves are coordinated.







The effects of correlation on financial statistics that measure risk illustrate the importance of the correlation assumption. In the case without correlation produces results similar to the case without any variability in gas prices. The case with no variability is analogous to debt leverage and aggravates risk. With positive correlation, the risks decline substantially as shown in the credit spread on subordinated debt column and the probability of a total loss on equity column. As with mean reversion, the table demonstrates that correlation is an important factor in measuring risk. Whether you ultimately attempt to make mathematical simulations or not, the exercises hopefully have demonstrated that thinking about whether an economic time series has mean reversion tendencies, boundaries, jumps or correlations with other variables is crucial in evaluating risk.

Chapter 23: The Difficult Problem of Estimating Volatility, Mean Reversion, Time Trends, Correlations and Price Boundaries from Historic Data and/or Market Data

You have probably noticed that none of the above discussion explained how to compute parameters for a time series equation. The exercises showed the importance of volatility, mean reversion, correlation and other variables, but not what the variables should be. Needless to say, if the construction of times series equations using parameters -- volatility, mean reversion, price boundaries, price jumps and correlations -- is to be useful in modelling, one must be able to calculate parameters for the models in an objective manner. This section discusses some of the practical and theoretical issues that arise when attempting to use statistical analysis of past data in computing various parameters. The discussion uses electricity and oil price data to demonstrate how the various calculations of volatility and mean reversion are made. Computing volatility parameters for a random walk process is described first. In subsequent sections, methods for computing volatility and mean reversion of a non-random walk process will be discussed.

The discussion focuses on parameters that should be input into a time series model rather than mechanical computation of statistics from historic data. To illustrate the difference, between input parameters and observed results, assume that a series of prices has a computed volatility of 20%, but that very tight upper and lower boundaries are modelled as part of the process. Because of the tight boundaries, the input parameter for volatility will not be the same as the resulting volatility in simulated cash flows.

Calculation of Volatility from a Random Walk Processes

Volatility is often defined as the standard deviation of annual percent changes in prices or some other variable. Since volatility is computed from the percent change rather than from absolute price levels, the unit of measurement for volatility is percentage – for example, volatility can be 20%, but it is would not be expressed as $30. Because volatility is expressed as an annual percentage rather than a daily or monthly percentage, if the standard deviation is computed from percentage changes in smaller time increments than annual increments, annualization adjustments are required. In the case of Brownian motion discussed below, for smaller increments than annual increments, the standard deviation is multiplied by the square root of the time increment.[4]

The three step procedure for calculating volatility of a random walk series without mean reversion involves first computing the rate of change in prices for the historic data. Second, the standard deviation of the series of rates of change is calculated. Finally, an adjustment is made for cases in which the rate of change is computed in time increments different from the time dimension of volatility (e.g. smaller than annual time increments.) The following formulas show how to compute volatility using both discrete and continual compounding with the three step process.[5] In both the discrete and continual cases, the first step is computing the rate of return over the period for reported prices. For continual compounding the rate of return is computed using the natural log:

$$\textit{Rate of Return}_i = \text{Natural Log } (\text{Price}_i/\text{Price}_{i-1})$$

Once a series of rates of return are established, standard deviation of the periodic rate of return is measured. If the prices are reported on an annual basis, the standard deviation of annual returns is volatility.

$$\textit{Period Volatility} = \text{Standard Deviation } (\textit{Rate of Return}_i)$$

The final part of the process for computing volatility where periodic prices are not expressed on an annual basis is converting the periodic volatility to an annual figure. Because of the mathematical process that defines Brownian motion, standard deviation of the rate of return increases with longer time periods. The variance of Brownian motion increases directly with time and the standard deviation increases with the square root of time. This means the period volatility defined above is multiplied by the square root of time measured in years (τ) to develop the annual volatility. Annual volatility is therefore defined as:

$$\textit{Annual Volatility} = \text{Standard Deviation } (\textit{Rate of Return}_i) \times (\tau)^{1/2}$$

There are a variety of ways to compute the volatility of a time series that does not have mean reversion price boundaries or price jumps. One can compute volatility directly from historic price changes; one can compute the standard deviation of historic price changes divided by the average price; or one can use regression analysis of the change in price against the prior period price. Alternatively, if traded options exist for the variable in question, one can compute the volatility parameter that the market

[4] For an explanation of why the standard deviation is multiplied by the square root of the time increment, see Pindyck, Robert and Dixit, Avinash, *Investment Under Uncertainty*, Princeton University Press, 1994, page 71.

[5] Section 2 of the workbook describes how to create a user defined function in spreadsheets that compute volatility whereby on can simply find the volatility of a series as one would find the average of a series.

believes will occur in the future. The different measures of volatility for the S&P500 using different time periods are shown below:

In a random walk (Brownian motion) time series, the length of the time period used to measure volatility does not significantly impact the estimated volatility parameter. Computation of volatility is confirmed by inputting volatility into a Monte Carlo simulation and computing the standard deviation of the resulting distribution divided by the mean price and by the square root of the time increment. The standard deviation grows over time, but the standard deviation divided by the square root of time remains at the volatility level input.

A practical issue that often arises in analysis of time series is how much historical data should be used in establishing parameters of the equation. A general rule in statistics is that if the structure of the economic variables has not changed, as much data should be used as possible to maximize the number of degrees of freedom. If the structure of the market has changed, then data should be used since the change in structure. Of course, the problem is judging when a structural change resulting from new merchant capacity, changes in demand or different fuel prices has occurred of sufficient magnitude to warrant truncating the data set. The problem in resolving this data issue is one of the problems with relying solely on time series analysis in projecting future prices.

Volatility of annual electricity prices computed by measuring the standard deviation of annual rates of change is shown on the graph below. Volatility ranges from 13% for the UK to 48% for the State of Victoria in Australia. The volatility in the California market including the high prices in 2000 and 2001 was 101.6%. By contrast, the annual volatility in real oil and natural gas prices over the period 1976 to 2004 was 26.7% and 20.2% respectively.

The annual volatility must be consistent with the time series equations defined above meaning that over a large number of observations, volatility computed from the outcome of the model should be the same as the input. For example, if a variable does not have mean reversion, price spikes or price boundaries, and a volatility parameter of 20% is input into a model, the standard deviation of prices that are observed from historic data or modelled in a simulation should also be 20%. In the case of a random walk, the standard deviation of prices increases over time and the volatility parameter must reflect this fact. Volatility of a random walk time series can be estimated by computing the standard deviation of the percent change in price from one period to the next and adjusting the result for the length of the time period in historical data and the pricing model. Adjusting the standard deviation for the time increment is necessary because volatility is generally measured in annual terms and the standard deviation increases as more time passes in a random walk process.

Non-Constant High Volatility of Electricity

S&P 500 Volatility: 1994-2002				
Period		Standard Deviation of Rate of Return	Periods per Year	Annual Volatility Percent
Daily		1.19%	250	18.77%
Monthly		3.89%	12	13.47%
Annual		17.10%	1	17.10%

To demonstrate volatility properties of financial securities as compared to electricity prices, the S&P 500 index and New England electricity prices are evaluated. Volatility is computed in which returns are evaluated for prices in different time increments for example, daily prices, monthly average prices and annual average prices. Daily volatility is also computed for different periods such as different years, monthly periods within a year and semi-annual periods within a year. The tables below demonstrate that when various different periods and/or time increments are used for the S&P 500, the volatility if quite stable. On the other hand when the same analysis is done for electricity prices, the volatility varies dramatically. Data for the volatility analysis is included in the reference files described in Section 5 of the workbook.

S&P 500 Volatility from Daily Returns			
Period	Standard Deviation of Daily Rate of Return	Periods per Year	Annual Volatility Percent
1997	1.14%	250	18.10%
1998	1.28%	250	20.30%
1999	1.14%	250	17.98%
2000	1.40%	250	22.14%
2001	1.36%	250	21.48%
2002	1.65%	250	26.09%
2001-First Half	1.45%	250	22.88%
2001-Second Half	1.21%	250	19.20%
2001-First Month	1.55%	250	24.49%
2001-Second Month	1.07%	250	16.91%
2001-Third Month	1.83%	250	29.00%
2001-Fourth Month	1.01%	250	15.98%
2001-Fifth Month	1.09%	250	17.31%
2001-Sixth Month	0.86%	250	13.61%

The first table demonstrates how standard deviation of the rate of return is converted to volatility for the S&P 500 and that even though the monthly volatility is lower than the daily volatility, the statistics are fairly stable. The second table demonstrates that if different time periods are used in evaluating volatility from daily returns, the statistics remain similar in the 20% range.

In a time series process with very high levels of volatility and mean reversion such as electricity, the volatility measured in short-term periods is much higher than in long term periods. The fact that volatility for electricity decreases with longer time frames of measurement has important implications on the structure of an option. If the strike provisions are based on monthly instead of hourly prices the volatility and therefore the value of an option are reduced significantly.

Using daily prices, annualized volatilities of electricity can approach 500%. Obviously, the annual change in electricity prices from one year to the next is not 500% (except in California.) The measures of short-term volatility for electricity are dramatically higher than estimated volatility for other commodities and financial securities. For example, the volatility of the S&P 500 index is about 10%, while the volatility of natural gas is about 50% and the volatility of crude oil is about 30%.[6] The insert compares the volatility of electricity to other components.

Attempting to Measure the Presence of Mean Reversion in Historic Data

[6]Source: NYMEX presentation on futures contracts in 1995.

In the graph of natural gas prices above it is pretty obvious that mean reversion exists. The question is more challenging for other time series that seem to follow a wandering pattern. One way to test for mean reversion is to make a regression equation of the change in price relative to the last period price. If there is no mean reversion, the coefficient on the change in prices should be close to zero while if there is mean reversion, the coefficient should be significantly different from zero.

$$\text{Change in Price} = (P_t - P_{t-1}) = \alpha + \beta \times P_{t-1}$$

While there are some statistical problems with the β parameter which is biased towards zero (meaning that mean reversion is difficult to detect) you can easily compute statistics using the SLOPE, INTERCEPT and TTEST functions in excel. However, a reasonable estimate can be obtained and, if the parameter is significantly different from zero (i.e., the t statistic on β is above 2.0), the process is clearly mean reverting. In running the regression, there is no need to make fancy adjustments for autocorrelation or other complexities.

If the regression equation is estimated using annual data, estimated coefficients from the regression equation can be used to compute both the mean reversion parameter and the volatility parameter. If the data is expressed in monthly, daily or hourly terms, adjustments must be made to convert the regression estimates to annual parameters. Before discussing how to use regression estimates of α, β and the standard error of the regression to establish the mean reversion and the volatility parameter, expected parameters of the regression equation are described.

Consider first the case where the β coefficient is equal to zero. In this case, the change in price has no relationship to the last period price. If the change in price is independent of historic prices, the equation meets a basic presumption of Brownian motion and a random walk. In a random walk, changes in price are not a function of past prices. When the β parameter is zero, the standard error of the regression is about the same as the standard deviation of the change in price – there is no variation in the constant term and the prior price drops out of the equation. This means that for an annual data series, the standard deviation of the change in price divided by the average price is the volatility. If the β parameter is significantly different from zero, the regression suggests that mean reversion is present in the time series. Here, in the case with mean reversion, the price change depends to a certain extent on the history of prices.

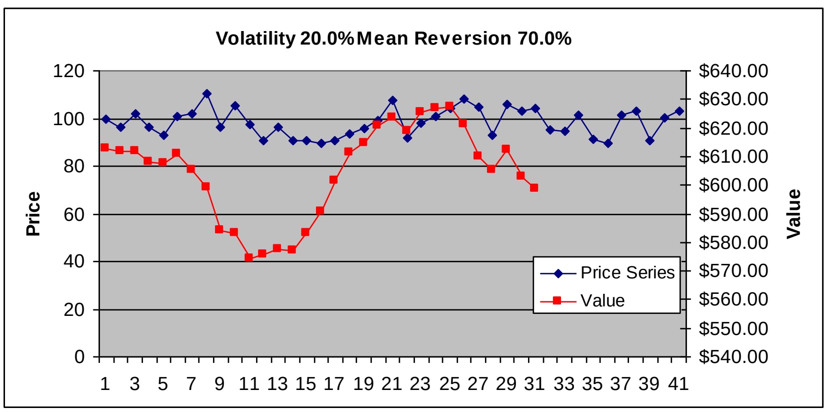
From a statistical standpoint, the mean reversion in the above graph can be demonstrated by computing the standard deviation of percent change in price for average prices in different time periods. With a long enough time series you could compute the average for one year, the moving average for three years, and the moving average for six years. Once these calculations have been made, the standard deviation of the percent change in three year and six year average prices can be compared to the standard deviation of price changes computed on annual monthly basis. Without mean reversion, the standard deviation of percent changes for the longer three year period should be six times as great as the monthly standard deviation (the square root of 3 x 12.) In fact, the standard deviation of percent changes for the longer period three year period is smaller than the standard deviation of the monthly percent changes as shown in the table below. The changing volatility is not present in stock prices where the volatility is approximately constant using weekly, monthly, annual or multi-annual periods.

Natural Gas Percent Changes

	Montly	Annual	3-Year
Standard Deviation	14.93%	28.29%	9.25%
Adjusted	51.71%	28.29%	5.34%

Measuring Mean Reversion, Volatility Futures and Asset Valuation

While prices themselves follow a mean reverting process, this does not necessarily mean that securities which are derived from the present value of future mean reverting processes are also mean reverting. Consider the example of a share in a company the produces oil where the major uncertainty in cash flow comes from oil prices. Assume also that the oil price itself follows a mean reverting process. In this case, even though the oil price is mean reverting, movements in the share value follows a random walk process with no mean reversion. The random walk process occurs because the expectation of a mean reverting process is already priced in the current share value. Anytime there is a random shock in the oil price, the share value moves, but the percentage change in the share value is not as large as the percent change in the oil price. The relationship between price movement and value movement is demonstrated in the graph below (the source of this graph is the Monte Carlo exercise from section 2 of the workbook).



A similar phenomenon occurs in evaluating the characteristics of the prices of forward contracts and the price of the commodity underlying the forward contract. Even though the price of the commodity may follow a mean reverting process, the movement in the value of the forward contract over time has less mean reversion. If the forward contract only has term of one month, then the time series characteristics of the contract are similar to the price of the commodity itself. In this case, if there is mean reversion for the commodity, there is also mean reversion for the future contract. However, if the contract is for a longer term such as ten years, then moves in the commodity price are moderated because of the mean reversion in prices as was the case with the share value example above.

Discuss example with oil prices.

Method 1: Construction of historic volatility
- How long
- Where to get the data
- Adjustments for time period
- Implicit assumptions – that the rate of return follows a normal distribution

Method 2: Computation of implied volatility
- Find option
- Plug in values including strike price
- Use appropriate model
- Use goal seek with macro to find volatility

Relying on implied volatility alone is risky. Implied volatility simply tells you how options are currently priced, but not whether they are realistically priced. Historic volatility, on the other hand, can help you understand whether or not options are currently cheap or expensive.

If make simulation, see if the simulated volatility of each series is the same as the expected volatility. This falls apart when there is mean reversion, price boundaries and price jumps.

Computation of Volatility and Correlation Time Series Parameters in Presence of Mean Reversion and Price Boundaries

Time period is crucial. Cannot apply monthly volatility to an annual model.

Computations of volatility are fairly straightforward if there is no mean reversion in the series. Where mean reversion is present, on the other hand, the calculations become more tricky. The first problem is that the mean reversion parameter measuring how long it takes for a series to come back to average levels is very difficult to extract from historic data. The second problem is that when volatility is computed from historic data and then used in a time series equation with mean reversion, the parameter that is input will not be the same as the resulting volatility output. The example below shows resulting volatility in a simulation where the input parameter for volatility is 20%. In the first case there is no mean reversion and in the second case mean reversion of 100% is assumed. The results are illustrated in the graphs below.

The general point of computing time series parameters is to derive statistics from historic data and then apply the statistics to mathematical models. If computation of a time series parameter from historic data is inconsistent with formulation of the mathematical model, this process will not work and the model will not produce unbiased results. For example, if 20% volatility is input into the model, on average, there should be 20% volatility in the simulated time series data that results from application of the model. With mean reversion and/or jumps and/or boundaries in models, the volatility formulas described above for a random walk series will not meet this proposition.

While the example produces the same results for both a mean reverting and a non-mean reverting series when the current price is the average price, the result does not hold when prices are different from the average. Consider a situation where the average price is 50, and the current price is still 100. In this case, a model driven by 20% volatility will produce different outcomes if the series does or does not include mean reversion. For the mean reversion series, the observed volatility from the 50 base price will be greater than 20% because the change in price includes the movement from 50 to 100 (a change of 100%) as well as effect of the 20% volatility. The time series without mean reversion will not have the added observed volatility driven by movements towards the mean and it will still have a 20%

volatility. Therefore, in the presence of mean reversion, if one measures volatility from the outcome of the prices using the standard deviation of price changes, the measured volatility will be more than 20% even though the volatility parameter driving the process is 20%. From the perspective of measuring volatility using observed historic prices, this means a computed volatility (say 100%) may be result from a time series process that has a lower true parameter driving the process (say 20%). In sum, if there is mean reversion, the volatility estimated from computing the standard deviation of the actual price changes will overstate the parameter required for the model.

In introducing mean reversion it was noted that computed volatility is not constant when measured over different time periods. Specifically, the volatility is greater when measured over short time periods than over long time periods. To demonstrate these problems consider a series with 100% mean reversion. Here, random movements away from the average are followed by movements back to the average level in the next period. Because prices keep coming back to the average level, the standard deviation of price changes does not increase over time as with a random walk series where prices wander indefinitely. This means that for mean reverting time series, one cannot easily estimate parameters using data with small time increments – days or months – and then derive annual parameters. If a time series computed on a daily basis has high volatility (e.g. daily temperatures) the daily volatility cannot be used to derive annual volatility (e.g. the volatility of the annual average temperature is much lower than daily volatility).

The problems of measuring volatility arise even where time increments of the data are the same as time increments in the model – for example where annual data is used to estimate parameters and the model simulates annual movements in price. Say that the underlying process for a mean reverting series and a random walk series without mean reversion both are driven by 20% annual volatility and that the current price is 100. If the current price is also the average price, both cases would produce similar results for the subsequent year driven by the volatility parameter (there is a 68% chance that prices in the next period are between 80 and 120.) In this case, the observed volatility measured as the standard deviation of the percent change in price is the same as the volatility parameter input into the model for both the mean reverting series and the non-mean reverting time series. For both cases, after the 100 starting point, the next period moves up and down in a similar manner.

If one knew the mean reversion parameter beforehand, one could first compute the expected price with the effect of mean reversion. In the example above with mean reversion of 100%, volatility could be computed as the standard deviation between the average price and the realized price. Using the example above where the price is 50, the volatility is computed as the standard deviation between the average price of 100 rather than the last period price. The problem with this approach is that the mean reversion parameter is not known beforehand.

It has been suggested that mean reversion parameters can be computed from a regression equation. This approach however produces weak statistical results and is not consistent with simulations. Given that it is somewhat complex.

Construction of Time Series Models from Supply and Demand Models

If an estimate of long-run prices is required to assess the risks associated with an investment that may last from twenty to forty years, annual time series parameters constructed from historic data are likely to have little relevance. The volatility estimates for annual price data will be very different from the shorter

term volatility parameters because, over the course of a year, weather reverts to normal levels, seasonality of maintenance schedules become less significant and other factors revert back to normal levels. Challenges arise in estimating annual volatility and particularly mean reversion for electricity prices because of the lack of historic price data for long periods and because there is less stability in the structure of the market (e.g. increases in capacity) the longer the time series. The next two chapters describe how long-run volatility can synthetically be computed using simulated data from supply and demand models. With supply and demand models, volatility is simulated from variation in loads, hydro conditions, fuel prices and maintenance outages.

Estimation of Other Parameters

In developing a time series model, a number of parameters other than volatility and mean reversion must often be estimated. Other parameters include time trends, equilibrium prices, lower and upper boundaries, correlations among variables, price jumps and jump arrival rates as well as volatility of price boundaries, trends and equilibrium prices. With the exception of correlations, there is little objective historic data that can be used to estimate the parameters. Instead, judgmental considerations must be the basis for the parameters. For example, in determining lower boundaries, one must have knowledge of the production function of suppliers in the industry. Similarly, the long-term marginal cost requires knowledge of the cost of constructing different types of new facilities and the existing supply situation.

Parameters for simulating a price jump can be estimated from historic data if it is assumed that there will be no changes in the structure of the market (e.g. no capacity additions, demand growth, changes in maintenance schedules.) Modelling a jump process required four parameters – (1) the probability of a jump, (2) the average size of a jump, (3) the standard deviation of a jump, and (4) the mean reversion of the jump. The following step by step process demonstrates an approach for incorporating price jumps into a time series equation derived from historic data:

> Step 1: Establish a level above which a price spike is defined. For example, assume that any price above $50/MWH is classified as a price spike.
>
> Step 2: Use the assumed price spike levels to compute the number of price spikes and the average amount of the price spike in the historic data – the number of price jumps divided by the total number of periods is the probability of the jump occurring.
>
> Step 3: Compute the average level of the price jump and the standard deviation of price spikes. These are computed simply computing statistics from the actual price spike data.
>
> Step 4: Evaluate the mean reversion of price spikes by measuring the time period before which price spikes disappear. By the definition of a price spike, the mean reversion of price spikes should be very high.

Appendix A:

Use of Continual Compounding Rather the Discrete Compounding in Time Series Equations

Finance and economic texts that describe Monte Carlo simulation and time series models often insist that it is proper construct simulations using continuous distributions rather than discrete distributions. If the assumption is made that returns come from a continuous distribution rather than assuming that changes occur at discrete intervals, some of the mathematical analysis becomes more confusing, but the basic ideas do not change. Those who insist on using continual distributions can be somewhat snobbish and scoff at models that assume discrete changes. The idea of this appendix is to work through differences between the continual and discrete equations so that you will not be intimidated by the differences.

To begin the discussion, consider the simple question of discrete versus continual compounding. The formulas for computing future values with different assumptions is shown in the table below. If the rate of return is relatively low (or the volatility is relatively low) the magnitude of differences is not very great.

Formulas for Continual and Discrete Compounding

Rate of Change	30%	
Beginning Value	100.000	
One Year - Discrete	130.000	=Beginning_Value*(1+Rate_of_Change)
One Year - Continuous	134.986	=Beginning_Value*EXP(Rate_of_Change)
Semi-Annual	132.250	=Beginning_Value*(1+Rate_of_Change/2)^2
Quarterly	133.547	=Beginning_Value*(1+Rate_of_Change/4)^4
Daily	134.969	=Beginning_Value*(1+Rate_of_Change/365)^365

If the rate of return is a large negative number, then the differences between continual and discrete compounding can be more dramatic. The table below shows that with a return of negative 300%, the discrete compounding case produces a negative number, while the continuous compounding case or the daily compounding case produces a small positive number.

Formulas for Continual and Discrete Compounding

Rate of Change	-300%	
Beginning Value	100.000	
One Year - Discrete	(200.000)	=Beginning_Value*(1+Rate_of_Change)
One Year - Continuous	4.979	=Beginning_Value*EXP(Rate_of_Change)
Semi-Annual	25.000	=Beginning_Value*(1+Rate_of_Change/2)^2
Quarterly	0.391	=Beginning_Value*(1+Rate_of_Change/4)^4
Daily	4.917	=Beginning_Value*(1+Rate_of_Change/365)^365

To understand why the continuous case produces a positive number, consider the case with daily compounding. In this case, the daily rate is -.82% meaning that on the first day of the year the 100 becomes 99.17. In the second period, this amount is reduces again and by the end of the year the base is reduced so that in the 364th day, the base is 4.958 resulting in the final number of 4.958 x (1-.82%) or 4.917. Because the returns are made smaller, the number continues to be positive. In time series analysis, the returns shown in the above table can be thought of as the volatility multiplied by the standard normal value. If the volatility is high – say 100%, and the normal draw is negative three, the outcome in the above example could occur.

The fact that prices cannot fall below zero with continual compounding is illustrated by the equation below where the percent change is computed from the natural log of the current price divided by the previous period price while the discrete change is computed from the standard formula (Current Price/Last Price – 1). In this equation, the term $e^{\text{Pct Change}}$ cannot be negative and therefore P_t also cannot be negative. The table below the formula demonstrates that when reductions occur, the measured percent change is above the discrete percent change – in absolute value terms. On the other hand, when prices are increasing, the measured continual rate of change is below the discrete measured change.

$$\text{Pct Change} = \text{Ln}(P_t / P_{t-1}), \text{ or } e^{\text{Pct Change}} = P_t / P_{t-1}, \text{ or } P_t = e^{\text{Pct Change}} \times P_{t-1}$$

Measured Percent with Discrete and Continuous Compounding

			Discrete	Continuous
Big Decline	Start	100		
	End	10	-90.00%	-230.26%
Increase			Discrete	Continuous
	Start	10		
	End	12	20.00%	18.23%
Decline			Discrete	Continuous
	Start	12		
	End	10	-16.67%	-18.23%

To see why the assumption of continual compounding makes time series equations a little more complicated, consider the case where the expected percent change in price is zero. If discrete equations are used, the likelihood of a lower price equals the likelihood of a higher price. For example, assume the volatility is 10%, meaning that the most likely percent change in price is zero, and there is a 68% change the price will increase or decrease by less than 10%. If the initial price is 100, and the percent change happens to be 10%, then the next price is 110. When the volatility is negative 10%, then the price is 90. The range in price on the upside and the downside is exactly the same. On the other hand, when continual compounding is used, then the expected price in the next period is not the same as the current price because the exponent of positive 10% does not result in the same price as the exponent of -10%. The table below illustrates the difference:

Base Price	100	
Volatility	10%	

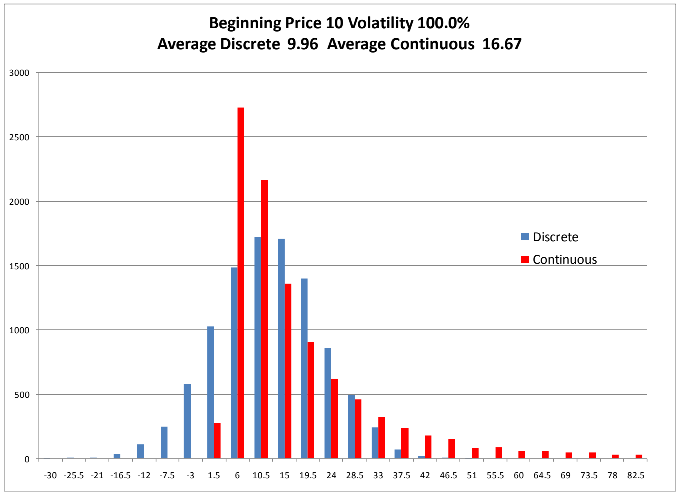
	Discrete			Continuous	
Sceanrio	Amount	Increase	Sceanrio	Amount	Increase
10%	110.00	10.00	10%	110.52	10.52
-10%	90.00	(10.00)	-10%	90.48	(9.52)
Average	100.00	0.00		100.50	0.50

While the small difference of .5 seems trivial, if the volatility is much higher, then the difference is much more as shown in the subsequent table which uses a volatility of 300% rather than 10%. Here the expected value difference is more than 900. To see why the difference is so large, return to the example of compounding with 365 days discussed above. Recall that the daily price change was .82%. When the price increased in the next period, then the .82% is applied to both the increase and the decrease. Similarly, if the price decreases, then the price increase or decrease is also applied. After going through 365 price increases or decreases, the low case is much less than the price increase case because the base of the price in the low case continues to decrease, while the base in the high case continues to increase.

Base Price	100	
Volatility	300%	

	Discrete			Continuous	
Sceanrio	Amount	Increase	Sceanrio	Amount	Increase
300%	400.00	300.00	300%	2,008.55	1,908.55
-300%	(200.00)	(300.00)	-300%	4.98	(95.02)
Average	100.00	-		1,006.77	906.77

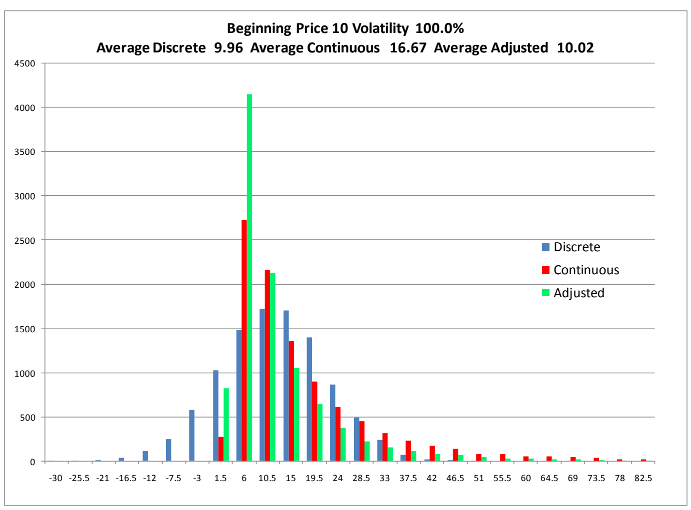
The above examples illustrated that the extent to which the average exceeds the initial price depends on the volatility. Further, the increase is non-linear. If 310% is used instead of 300%, the average increase is much more than .5 in the first table (it is more than 100.) A simulation can be easily constructed in excel to illustrate the effect of assuming a continual distribution versus a discrete distribution. Simply create a formula using the two alternative formulas – one with continual compounding using exp(volatility) and a second with discrete compounding using (1+volatility). Then multiply the volatility by a random draw from a normal distribution. To create the simulation, one can make a simple macro with a “for loop” and the “cells” function. A graph of price projections is shown below. The continual distribution cannot be negative and has a skewed distribution – a lognormal distribution -- while the discrete distribution has the same normal distribution as the percent change in prices.



In the above graph, the average price from the continual price assumption is 16.67 rather than 10, which is about the same as 10 x (e $^{(.5 \text{ x } (\text{vol}) \wedge 2)}$ -1). Specifically, the e $^{(.5 \text{ x } (\text{vol}) \wedge 2)}$ -1 is 6.49. This means that if we wanted to re-calibrate the expected price so that it still equalled the initial price, the term e $^{(.5 \text{ x } (\text{vol}) \wedge 2)}$ could be subtracted from each observation of the simulated price. If the general trend term is represented by α, then the time series equation with a time trend is:

$$P_t = P_{t-1} \text{ x } e^{(\alpha - .5 \text{ x volatility} \wedge 2 + \text{Standard Normal Draw x Volatility})}.$$

The adjustment affects the distribution of prices as well as the average as shown on the graph below. Note that after the adjustment, the average of the prices is very close to the original base price of 10.



Appendix B: Simulation of Correlated Variables

Many analyses require that the volatility and trends in more than one variable be incorporated in modelling analyses. For example, in evaluating the value at risk or the probability of default for projects sell or use include oil, natural gas and coal fuels, the risk of each of the variables must be analysed. This part of the chapter describes how the Monte Carlo process can be implemented when there are multiple variables.

When simulating the potential distribution of more than one variable, if the variables are not correlated with each other, the time series equations can be modelled as independent equations where the standard normal draws are independent from one another. To create a model with a series of independent variables, Monte Carlo analyses described above can be separately performed for each variable. However, if the variables to be simulated are correlated with one another, running the single Monte Carlo over and over will not produce accurate results. Instead, one must model the variables simultaneously using a analytical approach based on something known as Cholesky factors.[7] Modelling oil and gas prices that are input into a supply and demand model requires the use of this process.

Cholesky factors are derived from matrix algebra to demonstrate that in the case where two variables are correlated with one another, the standard normal draw for the correlated variable should be the weighted combination of a new random draw and the random draw for the first variable. To illustrate how Cholesky factors are used recall that Monte Carlo simulation is driven by applying a volatility parameter to a factor -- z_t -- that is driven by a random variable. For the oil price, one first makes a transformation of random variables using the standard normal variable resulting in values approximately ranging from -4 to 4:

z_1 = Inverse of Standard Normal(u_1 for first variable)

To apply a similar factor to the second variable – the natural gas price. The second standard normal draw is defined as z_2:

z_2 = Inverse of Standard Normal(u_2)

After the second random draw is made and filtered through the normal distribution, an adjusted factor to apply to the volatility is computed from the correlation between the variables. The adjusted factor is a weighted average of the first two factors according to the following equation:

$z\ adj_2 = z_1 \times \text{Correlation}_{1,2} + z_2 \times (1 - \text{Correlation}_{1,2}{}^2)^{1/2}$

The weighting of the two random draws depends on the correlation between the variables. Two extreme cases demonstrate the way correlation is implemented in a Monte Carlo simulation framework. If there is no correlation between the variables, the factor for the second variable should not be influenced by the random draw for the first variable. Therefore, with no correlation the entire weight of the standard normal draw is given to the second random draw as expected. On the other hand, if the variables are perfectly correlated, the entire process should be driven by the first variable and there should be no weight given to the second random draw. This again is expected because any random shock that affects the first variable should also affect the second variable.

[7] Jorion, Philippe, "Value at Risk", McGraw-Hill, 1997, pp. 242-243.

If three variables are correlated with each other instead of two variables, a similar process can be used. In this case, the correlation between the third variable and the second can be used in the equation and the equation does not require a memory of the random draw from the first variable. Since the second variable is already affected by the first variable, when the simulation process is finished, the first variable will be correlated with the third variable. Say the price of coal (variable 3) is correlated with the price of natural gas (variable 2) which in turn is correlated with the price of oil (variable 1). If the correlation between the price of gas and the price of coal is defined as $\rho_{2,3}$, the standard normal draw for the coal price is:

$$\text{Standard Normal Draw}_3 = \text{Standard Normal Draw}_2 \times \rho_{2,3} + \text{Inv Norm(New Random Draw)} \times (1 - \rho_{2,3}^2)^{1/2}$$

Here, the Standard Normal Draw_2 factor already incorporates the correlation between gas and oil prices.

Time Series Equations with Trend Terms and Movement to Long-term Marginal Cost

In financial analysis of long-term investments, the long-term trend in a variable is likely to be a more important consideration than the short-term fluctuations around the trend. Studying the long-term price trends through evaluation of long-term marginal costs and potential changes in productivity and evaluating demand analyses is essential in valuation analysis. The issue of developing price forecasts for valuation is addressed in other chapters. This section only describes how to incorporate long-term trends into time series analysis.

The simple random walk equation above can easily be extended to include trends in expected prices so that the expected price change is not zero. Trends can be incorporated in the equation through adding a drift term to current prices or through incorporating explicit forward price forecasts. The simplest way to add information about future expected directions in prices is to include a trend term to the random walk process. Differences in the equation that arise from an assumption of continual compounding is assumed is shown in appendix A where the trend factor is part of the exponent.

$$P_t = P_{t-1} + \alpha * P_{t-1} + \varepsilon.$$

The "α" term in the above equation is the percentage growth rate in prices and the ε term incorporates the volatility parameter and the distribution assumption as with the previous equation. The economic rational for adding a trend term in the equation could be inflation (if prices are to be forecast in nominal terms), expected capital gains from retaining profits (if the price is a stock price), or changes in industry productivity (if long-term real prices are projected.) If explicit forward forecasts are used, the α term varies by year.

Productivity is defined as the level of inputs needed to produce a level of output. For many things, one would expect technological improvements to result in gradual increases in productivity and reductions in the real prices over a long period of time because humans should find ways of using inputs somewhat more efficiently. Unless there are constraints on existing resources (such as oil and gas) or new regulatory requirements, the idea is that one can always employ the same technologies and management systems as those that exist today. If there are changes in management systems and/or technological improvements in the use of resources, these things improve productivity. With increased productivity in a competitive industry, prices should decline over time.

In modelling trend terms related to productivity, inflation, capital gains or other factors, it is reasonable to expect that the trend term itself is a stochastic variable (implying the growth rate variable is not known with certainty). If the trend term is stochastic, a volatility parameter related to the distribution of potential trend terms can be incorporated in the time series equation. Here, where a volatility parameter is included in the trend term, the trend term is modelled with the same process as the random walk equation above. Cases in which volatility is included both in the trend term and the current value can be termed a two factor process. To include a stochastic trend term, the trend term in the above equation is replaced with a trend term defined by a second random draw as shown in the formula below.

$$\alpha_t = \alpha_{t-1} + \text{Standard Normal}_1 \text{ x Trend Volatility x } \alpha_{t-1}$$

In the case of electricity prices, the volatility in short-term prices is driven by variability in weather, variation in hydro production and uncertainty in maintenance outages. On the other hand, volatility in the trend factor is driven by variation in future productivity that in turn is influenced by changes in the real cost of new capacity the heat rate of new plants. The volatility in short-term prices can be measured or simulated on an objective basis, but the volatility in trend factors is more difficult to estimate.

Price Jumps

Instead of limiting price movements with price boundaries, one may want to model sudden price movements that can occur in economic series. The review of stock prices and oil prices above demonstrated that the percent changes did not follow a normal distribution in part because there were occasional large movements that would not be occur in a normal distribution. The case study presented in Chapter 1 demonstrated the dramatic price movements that occurred in California electricity markets. Sudden movements can occur because of an event such as a war, a bankruptcy filing a strike, a new technology innovation, a large new market entrant or a political decision. The parameters required to model a jump process include the probability of the jump occurring, the size of the jump, the standard deviation of the jump when it does occur, and the mean reversion factor for the jump. The jump process is driven by the “arrival rate” or the probability that the jump occurs in the equation. Variables representing the arrival rate and the size of the jump are driven by random draws as for other stochastic variables. If a jump process is included in the equation, assumptions are required to model what happens after the jump has occurred. If a jump process is modelled, then one of the key assumptions is the speed of mean reversion over which the jump is eliminated. The size of the jump can be expressed as the volatility multiplied by the base value multiplied by a number from 3 to 6 to represent the number of standard deviations from the mean.

For electricity prices, the jumps caused by capacity constraints last over relatively short periods driven by heat waves and/or plant outages. These jump processes clearly revert back to a mean level and do not remain in place permanently (as may be the case with a technological innovation or a new market entrant). Modelling future price jumps depends on how the probability of jumps and the size of the jumps is affected by new capacity additions and demand growth. Intuitively, it is simple to measure the effect of price jumps on the value of a plant. If a power plant has a 30 year life and there is a 10% probability of a price jump arriving at any year of the plant’s life, there are three expected jumps over the life of the plant.

Care must be taken in modelling price jumps in a time series analysis. If the market is competitive, the effect of the price jump on profits may be quickly eliminated by competitive pressure and new companies coming into the market. If prices that are faced by an investment increase because of price

jumps, the effect is to increase returns rather than to increase the risks. Using the electricity case discussed throughout this section, a jump of four times the volatility is assumed to occur for .5% of the time. Even though the probability is low, the effect of the price jump is increase the possibility of earning high returns.

Appendix: Use of Regression Analysis to Compute Mean Reversion and Volatility Parameters

Estimation of Volatility and Mean Reversion from the Regression Equation

The objective of the statistical analysis discussed above is to ultimately develop a time series equation. This section describes how to convert coefficients of the regression equation into parameters of a time series model. The parameters include (1) the mean price (2) the mean reversion factor, and (3) the volatility.

Using the regression equation, the mean price is computed as:

$$\text{Mean Price} = -\alpha/\beta$$

The fact that this formula produces the mean price is demonstrated by simple algebra. Recall the formula for the change in price -- $\alpha + \beta \times P_{t-1}$. Over the sample period, the expected value of the change in price is zero. Further, the expected value of the last period price -- $E(P_{t-1})$ -- is the average price. Therefore, on an expected value basis, the formula reduces to:

$$0 = \alpha + \beta \times \text{Average Price},$$

that is rearranged to $-\alpha/\beta$ = Average Price.

The second parameter, the mean reversion factor, is computed from the β coefficient in the regression equation. Recall the mean reversion factor measures the average movement towards the mean price in a time period. This factor is defined as:

$$\text{Mean Reversion Factor} = -\log(1+\beta).$$

In this equation, if β is zero, the natural log of one minus zero is also zero, so the mean reversion parameter is zero. If β is -.63, the mean reversion factor is 1.0. If the β coefficient estimated from the regression equation is -.395, the mean reversion factor is .5 implying that in each period, prices move half way back to the mean level after a shock.

The third parameter of a time series equation that can be derived from the equation is the volatility. Volatility can be defined from the regression equation using the formula:

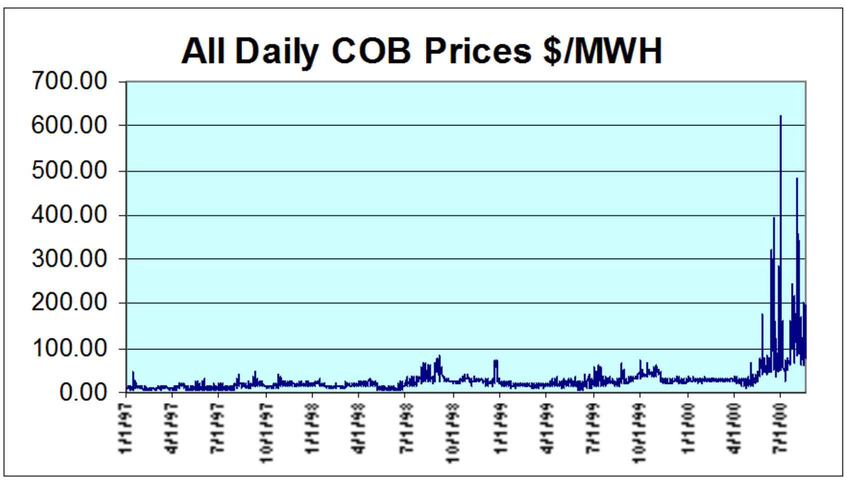
Absolute Volatility = Standard Error of Regression x $(\log(1+\beta)/((1+\beta)^2-1))^{(1/2)}$, and

Percent Volatility = Volatility/Average Price

In the above equation, if β is zero, the volatility level is the standard deviation of the change in price. The percent volatility is the standard error of the regression divided by the average price, or the standard deviation of the change in price divided by the average price. If β is greater than zero, the term $\log(1+\beta)$ is less than the term $(1+\beta)^2$, which means that the greater the β term, the smaller the volatility estimate. Intuitively, this means that if reversion to and back from the mean is causing some of the volatility, that volatility should be removed from the volatility that occurs exclusive of the mean reversion.

Calculating Mean Reversion Parameters

To illustrate calculation of parameters using the regression equation, this section reviews how a time series equation can be computed for prices at the California Oregon Border (COB) prior to the crisis in 2000. The graph below introduces the average daily on-peak and off-peak prices since 1997.



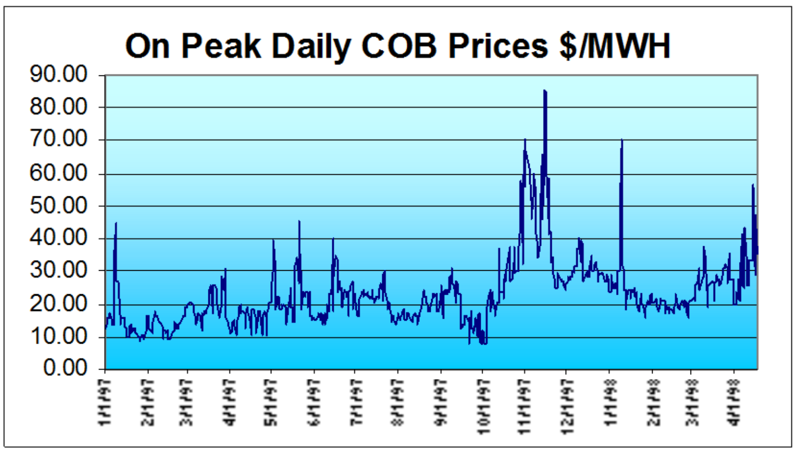
The accompanying graph demonstrates that price spikes began to occur in mid 2000, but before that, there were no prices above $100/MWH. To focus on computing parameters for the time series equations without the price spikes, only the data before 2000 is used in the statistical analysis.

While the formulas described above for the mean reversion factor, the mean prices and volatility seem straightforward, application to actual data raises a number of practical issues. Some of the issues involve when to separate models according to time period (summer versus non-summer and on-peak versus off-peak hours), how to account for price spikes and how much historical data to use in projections. Inspection of the data and knowledge of the markets clearly demonstrates that separate models should be developed for on-peak and off-peak prices. If one attempts to construct a model that includes both on-peak and off-peak prices, one will be trying to simulate a “saw tooth” pattern that will add nothing to the ability to make forecasts and will distort the statistics.

The graph above shows that in the summer of 2000, extended periods of very high prices occurred in both the on-peak and off-peak COB markets. These prices are a result of the much publicized power shortages in California that extended into 2001. Modelling the price levels after 2000 requires developing jump process parameters, which are discussed in a separate section. To focus the discussion on parameters that model price movements other than the price spikes, the remainder of this section concentrates on prices from 1997 through 1999. If there had been price spikes in the 1997-1999 data, we would have had to adjust the historic prices to subtract the price spikes.

Parameters for COB On-Peak Prices

The first series analysed for purpose of developing time series parameters focuses on on-peak prices through the end of 1999. These prices are shown on the graph below. The graph demonstrates that there is mean reversion in the data and the lower bound is around $10/MWH. Instead of prices wandering aimlessly after a shock, the prices clearly move back to an average level. Every time the prices move up significantly, the prices later move back.



Using the COB on-peak price data, the regression analysis of the change in price versus the last period price produces estimates shown in the accompanying table. The t-statistic on the "last_price" variable is above 6 in absolute value, which suggests that the previous price has a strong relationship with the price change. This implies that from a statistical perspective, mean reversion exits.

```
UNWEIGHTED LEAST SQUARES LINEAR REGRESSION OF CHANGE

                                PREDICTOR
VARIABLES    COEFFICIENT    STD ERROR     STUDENT'S T       P
---------    -----------    ---------     -----------    ------
CONSTANT        2.31568      0.37047          6.25       0.0000
LAST_PRIC      -0.09108      0.01334         -6.83       0.0000

R-SQUARED           0.0458      RESID. MEAN SQUARE (MSE)     23.4575
ADJUSTED R-SQUARED  0.0448      STANDARD DEVIATION           4.84329

SOURCE        DF       SS          MS         F        P
----------   ---   ----------   ----------   -----   ------
REGRESSION     1     1093.99      1093.99    46.64   0.0000
RESIDUAL     972     22800.7      23.4575
TOTAL        973     23894.7
```

Parameters for Time Series Equations Using Alternative Time Periods

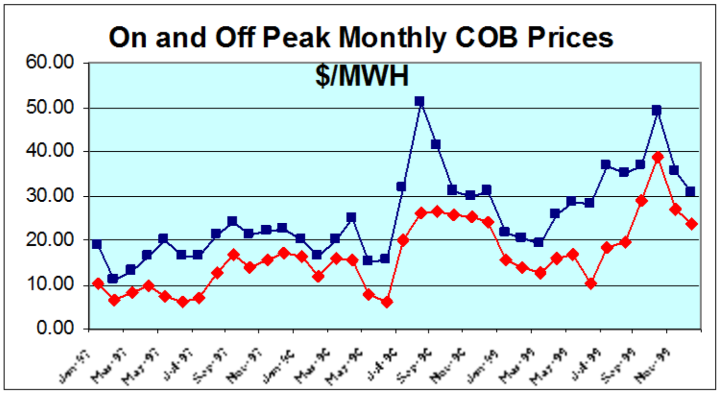
Estimated parameters differ depending on how the data is compiled. For example, monthly mean reversion differs from daily mean reversion and on peak mean reversion differs from off peak mean reversion. The volatility and the mean reversion parameters should be different for on-peak versus off-peak prices and for summer and non-summer prices. Parameters estimated from the regression equations for on-peak and off-peak time series are shown on the table below.[8]

	Off Peak Without Added Variables	Off Peak With Added Variables	On-Peak Summer	On-Peak Non-Summer
Regression Constant (a)	0.62404		3.28443	1.91129
Last Price Coefficient (b)	-0.03705	-0.07723	-0.11085	-0.08089
Standard Error of Regression	2.241	2.21711	6.60157	3.73169
Mean Price (-a/b)	16.84	16.84	16.84	16.84
Mean Reversion Factor (-log(1+b))	3.78%	8.04%	11.75%	8.43%
Multiplier (log(1+b)/((1+b)^2-1))^.5	0.720	0.736	0.749	0.737
Volatility	9.59%	9.68%	29.36%	16.33%
Annual Volatility	152%	153%	464%	258%

The table demonstrates that off-peak mean reversion and volatility is lower than for on-peak prices. Further, the volatility and mean reversion is less for the summer periods than for non-summer periods. This analysis suggests that separate models should be developed for alternative time periods. Section 2 of the workbook describes how to use spreadsheet techniques to efficiently compute volatility and mean reversion using spreadsheet techniques.

For some valuation issues, estimation of monthly price movements may be more relevant than projecting daily prices. For example, many purchase power contracts and tariffs have monthly prices rather than annual prices. As discussed above, for mean reverting time series, one cannot use data from daily data to project monthly parameters. Instead, monthly prices must be computed and then the parameters are computed from the monthly data. Monthly data for on-peak and off-peak COB prices are shown on the graph below.

[8] Annual volatility is presented using the square root of time as in the random walk method even though with mean reversion, the same time increment of the data should be used as the parameter for the time series model.



Regression analysis on the monthly prices demonstrates that the mean reversion and standard deviation are larger than with the daily prices. However, these larger numbers reflect a longer time period. Therefore, once prices are simulated, the deviation in prices may be more for the daily price series.

	Monthly On-Peak	Monthly Off-Peak
Regression Constant (a)	5.86202	2.8054
Last Price Coefficient (b)	-0.3429	-0.09676
Standard Error of Regression	6.55338	5.2408
Mean Price (-a/b)	17.10	28.99
Mean Reversion Factor (-log(1+b))	41.99%	10.18%
Multiplier (log(1+b)/((1+b)^2-1))^.5	0.860	0.743
Volatility	32.95%	13.44%
Annual Volatility	114%	47%

PARKED

The advantages and disadvantages of stochastic analysis in risk assessment are organized into five sections. After describing some general terminology associated with stochastic risk analysis, a simple example of process is presented. This simple example demonstrates how it is very easy to construct a time series equation and apply Monte Carlo simulation to compute statistics such as the probability of default, value at risk and distributions of returns. (You can complete a Monte Carlo simulation in excel in a matter of minutes.) After presenting this simple case, the discussion moves to subtle but important issues in constructing alternative types of time series equations that incorporate alternative volatility, mean reversion, correlation, price boundaries and price jumps in economic variables. The fourth section describes how to use historic data in measuring parameters for a time series model – volatility, mean reversion, price boundaries, correlations and jump processes. The final section describes how to apply time series models in making price forecasts with a Monte Carlo simulation.

Statistical concepts including time series equations and Monte Carlo simulation discussed in this chapter have been applied for decades in valuation of financial securities including equities, bonds, foreign exchange and other financial instruments. The famous Black-Scholes model that is so often used in all sorts of valuation applications assumes asset prices follow a stochastic process and come from a normal distribution. After the Black-Scholes formula became very widely accepted, the subject of applying mathematical analysis to real non-financial investments as well as financial investments gained popularity. There is now a lively debate as to whether stochastic analysis can be realistically applied in cash flow modelling for valuation. By working through various examples in this chapter, you can hopefully assess for yourself the efficacy of this technique.

Part of the attraction of using stochastic time series equations is to remove judgment from the process as described above. The controversy in over-relying on models and not using judgment is explained in the following statement by Alan Greenspan, a person who once had a lot of credibility and now may not be taken quite so seriously: “The essential problem is that our models – both risk models and econometric models – as complex as they have become – are still too simple to capture the full array of governing variables that drive global economic reality. A model, of necessity, is an abstraction from the full detail of the real world.”

One alternative to this is to combine judgment and with statistical analysis in creating time series equations. This sort of approach is suggested by a project finance analyst named Paul Ashley: “Given plausible assumptions for key risk drivers, the [credit] ratings are then built around simulating cash flows and debt coverage for individual projects. These simulations can then determine probability and loss distribution and directly observe the impact on cash flows and riskiness of a particular transaction structure under different market environments. The parameterization of the rating simulations combines market data and expert judgment in a structured way that allows transparency and consistency in risk evaluation without losing the ability to capture the specifics of a transaction structure.”

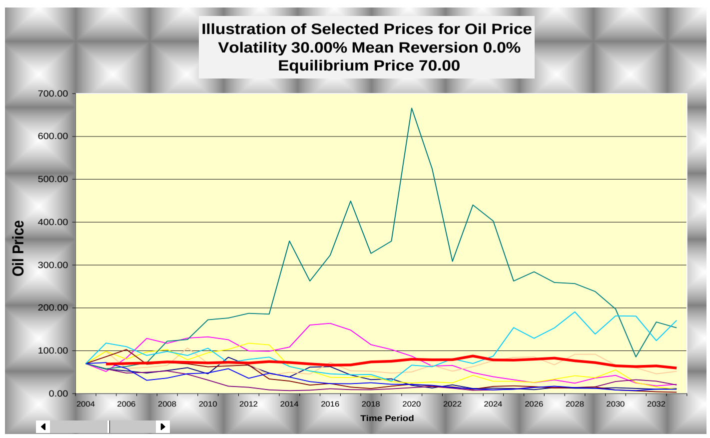
As with many statements made by financial professionals, the above statement requires a little translation. When the author uses the term simulation is used in the above quote, he means that some kind of time series equation is used together with Monte Carlo simulation. The word parameterization in the excerpt means converting predictions of key variables in a financial model into parameters such as volatility and mean reversion. The notion of transaction structure implies that once a Monte Carlo simulation has been created resulting in thousands of scenarios, this scenario data can be run with alternative debt levels, covenants, purchase prices and other factors to see how the simulation affects the probability of defaulting on debt (somewhat like running alternative transaction structures with scenario analysis discussed above.) Finally, the comment that expert judgment can be used means that it is possible to adjust time series parameters and to add factors such as lower and upper boundaries into an analysis which are derived from judgmental thinking about how low or high variables can move in the future.

Combining Judgment and Historic Statistical Analysis when using Probabilistic Equations to Represent a Business Venture

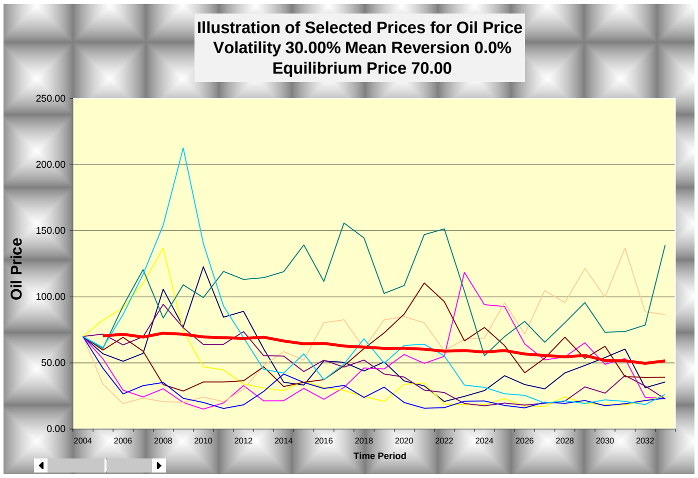
The manner in which judgment can be incorporated into stochastic analysis is somewhat analogous to the break-even discussion above. Recall that break-even analysis involves coming up with the

minimum or maximum level of a variable that allows the transaction to work. For example, one would come up with an oil price and then compare that price to historical trends, other forecasts, long-term productions costs and other factors. Similarly, the numbers produced by simulation should be inspected to see if they are consistent with historical observations of the variable and your judgment. A few specific examples of including judgment in a time series analysis include:

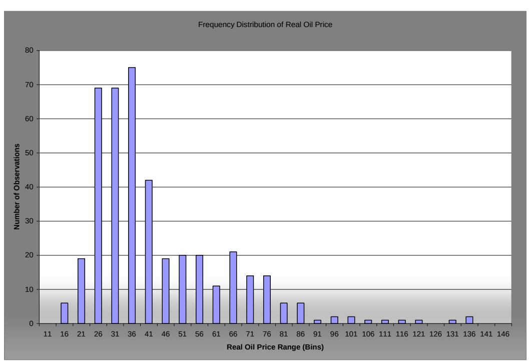
- Look at a number of a number of the scenarios for price, demand or other variables that are modelled with Monte Carlo simulation and evaluate the ranges from a judgmental perspective. For example, if many of the scenarios result in a price of zero or an unrealistically high price, the volatility, mean reversion and other parameters should be revised. The graph below shows selected scenarios that result from an assumption of no mean reversion, no upper or lower bounds and 30% volatility. Many of the scenarios result in a price of zero and some result in a price of more than 600. Clearly some judgment must be used to adjust the prices.

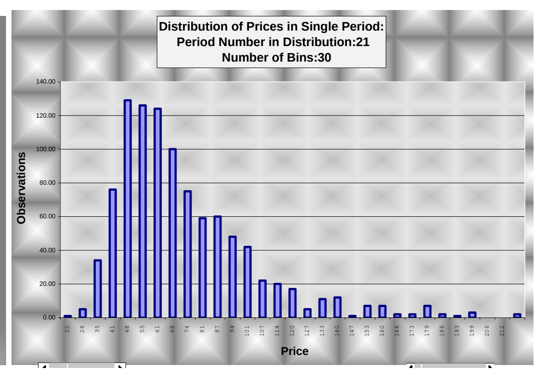


- Impose upper and lower boundaries on a judgmental basis. In the above example, one might estimate the short-run marginal cost to be 20 and the amount before which people stop consuming at 200. This judgmental adjustment of the parameters causes a dramatic difference in the simulations as shown in the graph below.



- Examine the frequency distribution of the variables that results from the Monte Carlo simulation and compare the simulated distribution with the actual historic distribution. If the simulated distribution has a much different appearance in terms of skewness and variation than the historical distribution, then you should go back and further study the time series parameters. The example below compares a simulated distribution with mean reversion to an actual distribution of oil prices. Simulated prices are shown on the chart on the right and the distribution of actual prices is shown on the left.





The mechanics of creating a frequency distribution shown in the above graph involve using the FREQUENCY function in excel. This function is called an array function because the output goes into multiple cells instead of a single cell. To accomplish this, first list the ranges in the variable that will be shown on the x-axis of the graph. These x-axis ranges are named bins. Once numbers for the bins are entered, shade the area next to the bins and then, while the range is shaded, type the function =FREQUENCY(data,bins). Arguments for the frequency function are first the data being evaluated and then the bins into which the data will be summarized. For example, if there are 70 observations of the oil

price between 35 and 40, then the result of the frequency function will be 75. This function will be used again below in evaluating whether numbers come from a normal distribution.

- Later in the chapter we will discuss including a jump process in time series equations. Adding this to a model may sound sophisticated, but it simply means that there is some probability (usually pretty small), that something really good or really bad will happen. The suggestion that these very unlikely events can be measured from statistical analysis and/or analysis of historic data is simply disingenuous. The whole idea of a sudden an very unexpected jump (like the effect of the Lehman Brothers collapse on the S&P 500) is that the event is unexpected and that it has not happened before. It is hard to imagine how a jump process can be added to time series models without some business judgment.

- Some of the judgmental techniques described above can be combined with stochastic modelling where one performs sensitivity, break-even or scenario analysis on parameters such as volatility and correlation to examine how sensitive risk analysis is to alternative parameters in time series models.

This one column model is in turn used as the basis for the time series equations. In order to construct a simulation with both senior and subordinated debt, the figure below shows how create a model which first sweeps all cash to pay senior debt, then, once senior debt is paid, a similar cash flow sweep is used to pay subordinated debt and finally, once subordinated debt is repaid, the remaining cash flow goes to equity holders. To make a model with these cash sweeps, the debt balance for senior and subordinated debt are set-up as described in the last chapter and cash flow sub-totals are computed to test the ability to pay-off debt securities. The repayment of debt is computed by using all available cash flow less interest expense until all of the debt is repaid (the minimum of the opening balance or the cash flow less interest.) Once the cash flow after senior debt is computed, the same process is used for subordinated debt. There is a small complication in the subordinated debt calculation because interest must be capitalized until the cash flow after senior debt is positive. If the financial model is set-up in this manner, then the closing balance of the senior and the subordinated debt measures the default on debt. In the example below, there is no default on senior debt and a default of 208 on subordinated debt.

The present value of the default debt provides the basis for computing the minimum credit spread. In the Using a risk free rate for the senior debt and the subordinated debt (the addition of credit spreads is addressed in the next chapter), the financial model below illustrates how defaults and equity cash flow can be computed.

Inputs

Operating	
Purchase Cost	1,000.00
Cash Flow	100.00
Volatility	20%
Financing	
Risk Free Rate	5%
Senior Debt	700.00
Subordinated Debt	200.00

Outputs				
MIRR to Equity	-100.00%	NPVo of Equity at Rf		($95.24)
MIRR to Project	2.32%	NPV of Project at Rf		($216.98)
	PV	Debt		FV
Default on Senior Debt	-	700.00		-
Default on Subordianted Debt	127.83	200.00		208.22

Model

Year	0	1	2	3	4	5	6	7	8	9	10
Cash Flow		100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00
Senior Debt											
Opening Balance	-	700.00	635.00	566.75	495.09	419.84	340.83	257.88	170.77	79.31	-
Less: Repayment	**-**	**65.00**	**68.25**	**71.66**	**75.25**	**79.01**	**82.96**	**87.11**	**91.46**	**79.31**	**-**
Add Debt Issued	700.00										
Closing Balance	700.00	635.00	566.75	495.09	419.84	340.83	257.88	170.77	79.31	-	-
Interest Rate	5%	5%	5%	5%	5%	5%	5%	5%	5%	5%	5%
Interest Expense	**-**	**35.00**	**31.75**	**28.34**	**24.75**	**20.99**	**17.04**	**12.89**	**8.54**	**3.97**	**-**
Cash Flow After Senior Debt	-	-	-	-	-	-	-	-	-	16.73	100.00
Subordinated Debt											
Opening Balance	-	200.00	210.00	220.50	231.53	243.10	255.26	268.02	281.42	295.49	293.54
Add: Interest Capitalised	-	10.00	10.50	11.03	11.58	12.16	12.76	13.40	14.07	-	-
Less: Repayment	**-**	**-**	**-**	**-**	**-**	**-**	**-**	**-**	**-**	**1.95**	**85.32**
Add: Debt Issued	200.00										
Closing Balance	200.00	210.00	220.50	231.53	243.10	255.26	268.02	281.42	295.49	293.54	208.22
Interest Rate	5%	5%	5%	5%	5%	5%	5%	5%	5%	5%	5%
Interest Expense	0.00	10.00	10.50	11.03	11.58	12.16	12.76	13.40	14.07	14.77	14.68
Interest Capitalised	0.00	10.00	10.50	11.03	11.58	12.16	12.76	13.40	14.07	0.00	0.00
Interest Paid	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**14.77**	**14.68**
Cash Flow to Equity	-100.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Free Cash Flow	-1,000.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00

ENRON attempted to profit from trading in markets with high volatility. The following excerpt from ENRON's marketing materials illustrates this strategy: Electricity is by Far the Most Volatile of Commodities: Recent crude oil prices have seen price fluctuations as high as 70 percent. The volatility of coal spot prices over the past year in Europe has averaged 20 percent. Nickel, the most volatile metal traded, has seen volatility of 60 percent, and over the past five years, the price of natural gas futures has varied by 100 percent. However, none of these volatility figures compare to power prices this year in Europe. This January on the Amsterdam Power Exchange, the price of electricity suddenly jumped from 44 euros/MWh to 474 euros/MWh - a volatility of about 1,500 percent. The biggest factor at play in power markets is supply and demand. But because consumers don't know, in real time, how much electricity they are using, they can't react to high prices by cutting their consumption. According to Andersen Consulting, "Wholesale power markets, during periodic supply/demand crunches, have exhibited volatility 20 to 30 times that seen in financial and oil markets. This is driven by the fact that power cannot be truly stored."

The three step process above assumes that the distribution of prices follow a normal distribution. This assumption that price movements follow a normal distribution is not necessary. In converting the probabilies of making a particular random draw to a simulate price series, Monte Carlo simulation can be used with any type of probability distribution deemed appropriate to represent the nature of a price moves. Filtering random draws through a normal distribution implies that it is more likely for the

variables to fall near the prior period price than far away from the prior period price. The graph below shows the standard normal distribution where the mean value is zero and the standard deviation is 1.0.

The instructions in the above paragraph involving copying the time series equation gets the Monte Carlo job done. The things that were done to create a simulation are described in a more formal way below. Applying Monte Carlo simulation to the equation above involves a few steps where a time series equation is defined, random numbers are generated, a path of prices is computed and the process is repeated for multiple scenarios. The following six step procedure describes the general process for creating a Monte Carlo simulation:

Step 1: Choose the time increment (the period could be a day, a month or a year), the number of periods for each path, and the number of simulated price paths.

Step 2: Begin the process with the current price and draw a random number between zero and one to project how prices in the first period will change from the first period to the second period.

Step 3: Convert the random number into a standard normal draw through filtering the random number through a normal or another distribution. This means using the inverse of the normal distribution as the normal probability distribution defines the probability of a number such as zero (the mean) and plus or minus one (the standard deviation). If the filter is the normal distribution, random number draws between 0 and 1 result in filtered numbers between -4 and 4 with a mean of zero and a standard deviation of 1.

Step 4: Use the filtered random number – the standard normal value -- to compute the next period price in the time series equation.

Step 5: Repeat the process of drawing a random number, filtering it through a probability distribution and computing a new price for subsequent periods in the first scenario.

Step 6: Once the process is completed for the total number of periods for one price path, begin the same process for the next simulation and continue to repeat the procedure for the total number of simulations.

To illustrate how this Monte Carlo simulation process works, assume a cash flow forecast is made for 10 years into the future with a volatility of 20%, a starting cash flow of 100 and no mean reversion or time trend. The simulation begins by making a random draw for the first month and the first cash flow scenario, say the draw results in a number of 0.6. Using the inverse of a standard normal distribution, the value of 0.6 translates into a standard normal value of .25 (a random number of 0.5 would result in a standard normal value of zero.) This standard normal value is then multiplied by the volatility of 20% yielding a percent change in cash flow of .25 x 20% or 5% from the first year to the second year. The cash flow in the second year therefore becomes 100 x (1+0.05) or 105. To model cash flow in the third year for this first scenario, the same process is used – drawing a random number, filtering it through an inverse normal distribution and adjusting for volatility -- but the cash flow that is the starting point for the formula is 105 instead of 100. After cash flow in the third year is established, another random draw will be made resulting in a new cash flow for the fourth period. By the 10th year, a price series is created that reflects the effects of making 9 random draws. Once the entire first cash flow scenario is established, the same process is used to generate a second set of cash flow for 10 years. Ultimately, the procedure for computing cash flow paths is repeated many times for the 10 year period – often making 1,000 to 100,000 price simulations. Once all the cash flow paths are simulated, the distribution

of cash flow from the 20th month can be analysed through computing the standard deviation and other statistics.
The volatility parameter used in the time series equation did not include parameters for mean reversion, trend factors, boundaries, jumps, long-run equilibrium values and correlations with other prices to project how variables evolve over time. Mean reversion refers to the tendency of a variable, after it increases above the mean, to move back to the mean. For example if the long-term mean of the oil price is $65 per barrel, and the current price is $100 per barrel, then the price may have a tendency to move back to the mean value. The speed at which a variable moves back toward the mean is called the mean reversion factor which can be added to a time series equation. Other factors can be developed for trend parameters, correlation parameters, lower boundaries, upper boundaries and sudden jumps. Details of how these parameters are included in time series equations and alternative methods for computing the parameters is described in subsequent sections.

In describing how various issues such as how mean reversion, price trends, correlations, price boundaries and price jumps affect time series analysis and ultimately risk measurement, a relatively simple model of an electricity plant is used to illustrate the effect of different parameters. The plant is assumed to earn revenues from an uncertain price and capacity factor and it has to purchase natural gas for generation. It is assumed to have a cost of 500 million with 200 senior debt, 50 of subordinated debt and 250 of equity. The initial exercise assumes the price of electricity comes from a Brownian Motion process with different volatility parameters of 15%, 25% and 40%. Variables such as the capacity factor and the natural gas price are assumed to have no volatility. This case is similar to the simple case discussed in the earlier section as there is no mean reversion, price boundaries or other time series parameters other than volatility (except that a log-normal distribution is assumed which is contrasted with a discrete distribution in Appendix A.)

The parameters required to run a simulation for a random walk include the starting price, the volatility, the number of periods, the number of simulations and whether the data is transformed to logs. To illustrate this process, the exercise assumes beginning price of 80, volatility of 15%, 25% and 40% and 20 annual periods for the analysis. 1,000 simulations have been run to achieve a distribution of possible prices.

The real world activities of developers compared to the theoretically rational behaviour assumed in many pricing models is made by Standard and Poor's as follows.[9]

> Generally most models assume rational behaviour by market participants, and that a system will operate economically in the most efficient manner, given the physical constraints of the system. Unfortunately, market participants do not always behave rationally or uniformly – fortunately they do not behave recklessly either…Underlying every model is the assumption that the markets will operate in a perfect long-term equilibrium, and that they will price electricity at the marginal cost of production. Unfortunately the reality of energy markets appears much different. Equilibrium lasts momentarily at best. A sudden change in fuel prices, a change in system load, or the presence of a new power station will upset the equilibrium.

[9] Rigby, Peter N., "Risks from Left Field: Is There a Problem with Pricing Models," Comments at the 15th Annual Global Power Markets Conference, April 16-18, 2000.

Problems with Time Series Equations for Analysis of Electricity Prices

This chapter introduced the characteristics of competitive electricity prices and it presented the mechanics of developing time series analysis. Before moving to other subjects, a few comments on problems with the times series analysis applicable to electricity are in order. Four issues and problems include:

(1) Time series analysis is more powerful when used together with supply and demand models than on a standalone basis in evaluating the risks associated with investment decisions;

(2) Regression analysis often does not produce adequate parameters for estimating time series equations when there is a high degree of mean reversion;

(3) Using statistical analysis of historic data in time series equations will produce erroneous analysis when structural changes have occurred in a market such as capacity additions or increased demand;

(4) When parameters are computed from price data without additional analysis, the underlying sources of volatility from price levels such as weather, hydro conditions, maintenance outages and economic activity can be ignored.

The statistical analysis used to develop time series parameters is not a simple matter of running a regression equation. Instead, the mean reversion parameter and the volatility must be tempered with judgment. Because of the limited ability to use objective statistical analysis in deriving parameters of the time series equations, analysts are often left trying alternative parameters and evaluating whether the outcome of price paths are reasonable.

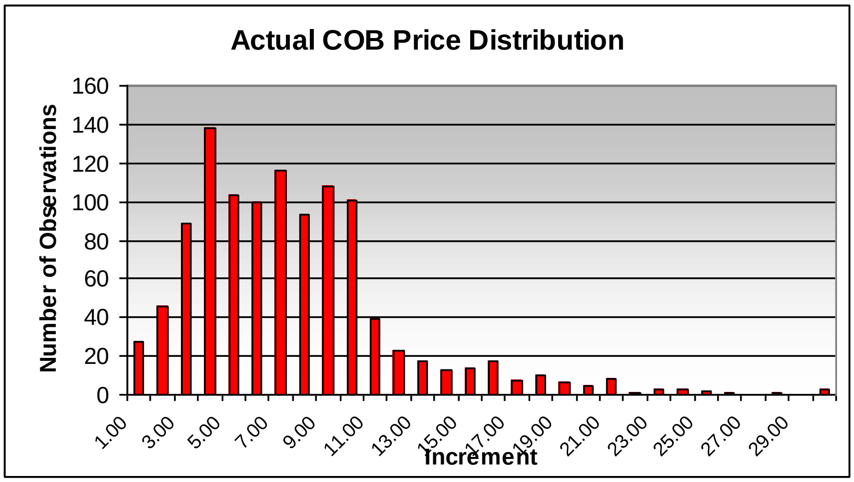
Adjusting parameters as described above is inadequate from an objective empirical standpoint. An alternative approach is to compare the distribution of simulated prices with the distribution of actual historic prices. The distribution of prices is computed by developing increments of price ranges and counting either simulated prices that fall within the range or by counting the number of actual occurrences that are in the range. An example of computing actual distributions is shown in the graph below. With the actual distribution computed one could test the simulation analysis with different parameters and compute the absolute difference between the simulated distribution and the absolute difference for each increment. This could be done on a relative basis as well as an actual basis as illustrated by the following formulas:

Actual Index = Actual Value/Average Value and Simulated Index = Simulated Value/Average Simulated

Using these index values, distributions are computed using the same bins for both the actual distribution and the simulated distribution as explained in section 2 of the workbook. Once these distributions are computed, the difference between actual and simulated distributions are computed and summed. The difference for each bin that is summed is:

$$\text{Difference}_{Bin} = \text{Absolute Value (Actual Index}_{Bin} - \text{Simulated Index}_{Bin})$$

Aside from the question of whether time series models adequately represent historic data, the models must be evaluated from a forecasting perspective. It is arguable that the massive financial problems that occurred from the demise of Long Term Capital Management arose from assumptions that the structure of markets – volatility and other parameters – could be estimated from historic data. If time series equations are entirely derived from historic data, the models cannot handle shifts in economic parameters resulting from changes in the structure of the market. For example, the addition of significant amounts of capacity in the market will reduce the probability of jumps and lower the volatility and mean price. Assessing how these parameters will change without delving into supply and demand models is virtually impossible.



A final problem with the time series analysis is that parameter estimation from actual prices ignores basic economic data on demand and supply drivers. For example, assume that it is known that prices were high because of abnormal weather or higher than normal plant outages. The analysis should not make prospective adjustments for normal conditions. Estimating parameters from actual prices does not do this.